

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2005, 19:46:59 ; Search time 234 Seconds  
(without alignments)  
2439.198 Million cell updates/sec

Title: US-09-855-828-1  
Perfect score: 4234  
Sequence: 1 MFKSLTKVKNVKPIGENNEN.....PSAEGGEEVLTIETVKEKAKQ 809

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot\_05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4234	100.0	809	1	Q9NQW8 homo sapien
2	3181.5	75.1	782	1	Q8MJJ7 canis famil
3	2515	59.4	694	1	Q9JJZ9 m cyclic-nu
4	1971.5	46.6	866	2	Q80XL8 mus musculu
5	1959	46.3	1339	2	Q35788 rattus norv
6	1952	46.1	1245	2	Q43636 homo sapien
7	1949	46.0	909	1	Q14028 homo sapien
8	1949	46.0	1251	2	Q9UMG2 homo sapien
9	1923	45.4	858	2	Q55157 rattus norv
10	1920	45.3	938	2	Q77658 BOVIN
11	1920	45.3	948	2	Q77659 BOVIN
12	1920	45.3	952	2	Q77660 BOVIN
13	1920	45.3	1394	1	CNGB1_BOVIN
14	1639.5	38.7	573	2	Q4SKK9 TETNG
15	1327	31.3	743	2	Q4T4Q5 TETNG
16	1258	29.7	624	2	Q4SKD5 TETNG
17	1174.5	27.7	429	2	Q4SKL1 TETNG
18	1160	27.4	910	2	Q5TQJ3 ANOGA
19	1103	26.1	1040	2	Q9W2D5 DROME
20	1096	25.9	800	2	P90975 CAEEL
21	1095	25.9	800	2	Q62237 CAEEL
22	1085.5	25.6	810	2	Q61DM6 CAEBR
23	878	20.7	515	2	Q95SL0 DROME
24	850	20.1	694	1	CNGB1_HUMAN
25	850	20.1	694	2	Q53RD2 HUMAN
26	842	19.9	676	2	Q4VAP7 HUMAN
27	831.5	19.6	683	1	CNGB1_RAT
28	829	19.6	663	1	CNGB2_BOVIN
29	826.5	19.5	698	2	Q6ZNA7 HUMAN
30	823.5	19.4	684	2	CNGB1_MOUSE
31	821.5	19.4	735	1	CNGB1_CHICK

32	818.5	19.3	664	1	CNGB2_RABIT
33	818	19.3	686	2	Q9N0H4_PIG
34	817	19.3	686	1	CNGB1_HUMAN
35	817	19.3	690	2	Q4W5E3_HUMAN
36	807.5	19.1	682	1	CNG ICTPU
37	806	19.0	611	2	Q9QWN7_RAT
38	806	19.0	632	2	Q9ER32_RAT
39	806	19.0	670	2	Q9ER33_RAT
40	804	19.0	664	1	CNGB2_RAT
41	804	19.0	664	2	Q549G7_RAT
42	801.5	18.9	691	1	CNGB1_CANFA
43	797	18.8	690	1	CNGB1_BOVIN
44	794	18.8	706	1	CNGB3_BOVIN
45	789.5	18.6	664	2	Q80XH6_MOUSE

ALIGNMENTS

RESULT 1  
CNGB3\_HUMAN  
ID CNGB3\_HUMAN STANDARD; PRT: 809 AA.  
AC Q9NQW8; Q9NRE9;  
DT 29-MAR-2004 (Rel. 43, Created)  
DT 29-MAR-2004 (Rel. 43, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Cyclic-nucleotide-gated cation channel beta 3 (CNG channel beta 3)  
DE Cyclic nucleotide gated channel beta 3) (Cone photoreceptor cGMP-gated channel beta subunit) (Cyclic nucleotide-gated cation channel modulatory subunit).  
GN Name=CNGB3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 1), TISSUE SPECIFICITY, VARIANT ACHM3 PHE-435, AND VARIANTS PRO-298 AND GLY-755.  
RX MEDLINE=20414632; PubMed=10958649; DOI=10.1093/hmg/9.14.2107;  
RA Kohl S., Baumann B., Broghammer M., Jaegle H., Slevings P., Keilner U., Spegal R., Anasatzi M., Zrenner E., Sharpe L.T., Wässinger B.;  
RT "Mutations in the CNGB3 gene encoding the beta-subunit of the cone photoreceptor cGMP-gated channel are responsible for achromatopsia (ACHM3) linked to chromosome 8Q21.";  
RL Hum. Mol. Genet. 9:2107-2116 (2000).  
RN [2]  
RP NUCLEOTIDE SEQUENCE OF 114-809 (ISOFORM 2), FUNCTION, SUBUNIT, VARIANT ACHM3 PHE-435, AND VARIANT CYS-234.  
RC TISSUE=Retina;  
RX MEDLINE=20347712; PubMed=1088875; DOI=10.1038/77162;  
RA Sundin O.H., Yang J.-M., Li Y., Zhu D., Hurd J.N., Mitchell T.N., Silva E.D., Maumenee I.H.;  
RT "Genetic basis of total colourblindness among the Pingelapese islanders.";  
RL Nat. Genet. 25:289-293 (2000).  
CC -1- FUNCTION: Visual signal transduction is mediated by a G-protein coupled cascade using cGMP as second messenger. This protein can be activated by cGMP which leads to an opening of the cation channel and thereby causing a depolarization of rod photoreceptors. Induced a flickering channel gating, weakened the outward rectification in the presence of extracellular calcium, increased sensitivity for L-cis diltiazem and enhanced the cAMP efflux of the channel when coexpressed with CNGB3 [By similarity]. Essential for the generation of light-evoked electrical responses in the rod, green- and blue sensitive cones.  
CC -1- SUBUNIT: Heterooligomeric complex with CNGB3.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q9NQW8-1; Sequence=Displayed;  
CC Name=2;

CC IsoId=Q9NQW8-2; Sequence=VSP\_009742;  
 CC Note=No experimental confirmation available;  
 CC -!- TISSUE SPECIFICITY: Expressed specifically in the retina.  
 CC -!- DISASE: Defects in CNGB3 are a cause of achromatopsia 3 (ACHM3)  
 CC [MIM:242300]; also known as Pingelapese blindness. ACHM3 is a  
 CC congenital complete achromatopsia and is distinct from total  
 CC colorblindness mainly because of the consistent concurrence of  
 CC severe myopia.  
 CC -!- SIMILARITY: Belongs to the cyclic nucleotide-gated cation channel  
 CC (TC 1.A.1.5) family.  
 CC -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; AF272900; AAF86274.1; -; mRNA.  
 CC EMBL; AF228520; AAF80179.1; ALT INIT; mRNA.  
 CC Ensembl; ENSG00000170289; Homo sapiens.  
 CC HGNC; HGNC:2153; CNGB3.  
 CC MIM; 605080; -;  
 CC MIM; 262300; -;  
 CC GO; GO:0007165; P:signal transduction; NAS.  
 CC GO; GO:0006810; P:transport; NAS.  
 CC GO; GO:0007601; P:visual perception; TAS.  
 CC InterPro; IPR000595; cNMP bd.  
 CC InterPro; IPR005821; Ion Trans.  
 CC InterPro; IPR001622; K+channel pore.  
 CC Pfam; PF00027; cNMP binding; 1.  
 CC Pfam; PF00520; Ion trans; 1.  
 CC SMART; SM00100; cNMP; 1.  
 CC PROSITE; PS00888; cNMP\_BINDING\_1; 1.  
 CC PROSITE; PS00889; cNMP\_BINDING\_2; 1.  
 CC PROSITE; PS00442; cNMP\_BINDING\_3; 1.  
 CC KW Alternative splicing; CGMP; cGMP-binding; Disease mutation;  
 KW Glycoprotein; Ion transport; Ionic channel; Multigene family;  
 KW Nucleotide-binding; Polymorphism; Sensory transduction; Transmembrane;  
 KW Transport; Vision.  
 FT TOPO\_DOM 1 216 Cytoplasmic (Potential).  
 FT TRANSMEM 217 237 H1 (Potential).  
 FT TOPO\_DOM 238 250 Extracellular (Potential).  
 FT TRANSMEM 251 271 H2 (Potential).  
 FT TOPO\_DOM 272 302 Cytoplasmic (Potential).  
 FT TRANSMEM 303 323 H3 (Potential).  
 FT TOPO\_DOM 324 359 Extracellular (Potential).  
 FT TRANSMEM 360 380 H4 (Potential).  
 FT TOPO\_DOM 381 417 Cytoplasmic (Potential).  
 FT TRANSMEM 418 438 Extracellular (Potential).  
 FT TOPO\_DOM 439 504 Extracellular (Potential).  
 FT TRANSMEM 505 525 H6 (Potential).  
 FT TOPO\_DOM 526 809 Cytoplasmic (Potential).  
 FT NP\_BIND 532 676 cGMP (By similarity).  
 FT BINDING 592 592 cGMP (By similarity).  
 FT BINDING 604 604 cGMP (By similarity).  
 FT CARBOHYD 468 468 N-linked (GlcNAc...) (Potential).  
 FT VARSPLIC 590 594 Missing (in isoform 2).  
 FT VARIANT 234 234 W -> C (in dbSNP:6471482).  
 FT VARIANT 298 298 T -> P (in dbSNP:4961206).  
 FT VARIANT 435 435 S -> F (in ACHM3).  
 FT VARIANT 755 755 E -> G (in dbSNP:3735972).  
 FT SEQUENCE 809 AA; 92250 MW; AC23B072C1C7DB3 CRC64;  
 Query Match 100.0%; Score 4234; DB 1; Length 809;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-238;  
 Matches 809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFKSLTKVKNVKPIGENNENEOSSRRNEBEGSHPSNQSOQTAAEENKGEKSLTKSTPV 60  
 DB 1 MFKSLTKVKNVKPIGENNENEOSSRRNEBEGSHPSNQSOQTAAEENKGEKSLTKSTPV 60  
 QY 61 TSEBPTHNIQDKLSKKNSSGDLTTNPDPQNAEAETGTVPEQKEMDPGKEGPNSPONKPPA 120  
 DB 61 TSEBPTHNIQDKLSKKNSSGDLTTNPDPQNAEAETGTVPEQKEMDPGKEGPNSPONKPPA 120  
 QY 121 APVINEYADAQHLNVLKRMORTALYKKLVGDDLSSEASPOAKPTAVPVKESDDKP 180  
 DB 121 APVINEYADAQHLNVLKRMORTALYKKLVGDDLSSEASPOAKPTAVPVKESDDKP 180  
 QY 181 TEHYRLLRFKVKKMPLEYKRIKLPSIDSYTDRLYLLMLLVTLAYNNWCWFIPLRL 240  
 DB 181 TEHYRLLRFKVKKMPLEYKRIKLPSIDSYTDRLYLLMLLVTLAYNNWCWFIPLRL 240  
 QY 241 VPPYOTADNIHYWLIADIICDIIYLYDMLFTQPLQFVRGGDIIVDSNELRKHYRTSTKF 300  
 DB 241 VPPYOTADNIHYWLIADIICDIIYLYDMLFTQPLQFVRGGDIIVDSNELRKHYRTSTKF 300  
 QY 301 QLDVASIIPFDICYLFFGFNPMFRANRMLKYTSFFEFNHLESIMDKAYIVIRTTGYL 360  
 DB 301 QLDVASIIPFDICYLFFGFNPMFRANRMLKYTSFFEFNHLESIMDKAYIVIRTTGYL 360  
 QY 361 LFIILHINACVYVWASNYEGIGTRVWYDGEQNEYLRCYYWAVRTLITIGGLPEPOTLFEI 420  
 DB 361 LFIILHINACVYVWASNYEGIGTRVWYDGEQNEYLRCYYWAVRTLITIGGLPEPOTLFEI 420  
 QY 421 VFQLLNFFSGVVFVSSLIQGMRDVIGAATANQNYFRACMDDTIAYNNYSIPKLQKVR 480  
 DB 421 VFQLLNFFSGVVFVSSLIQGMRDVIGAATANQNYFRACMDDTIAYNNYSIPKLQKVR 480  
 QY 481 TWYEYTWDSQRMLEDSDLLKTLPTTVQALALADNVNFSISKVDLPKGDOTMIYDMLLRL 540  
 DB 481 TWYEYTWDSQRMLEDSDLLKTLPTTVQALALADNVNFSISKVDLPKGDOTMIYDMLLRL 540  
 QY 541 KSVLYLPDGFVCKGEGIKEMWIIKHGEVQVLGGPDGTVKLVTLKAGSVFGESISLLAAGG 600  
 DB 541 KSVLYLPDGFVCKGEGIKEMWIIKHGEVQVLGGPDGTVKLVTLKAGSVFGESISLLAAGG 600  
 QY 601 GNRRTANVAHGFANLLTLDKKTLOEILVHVYPDSERIIMKKARVLLKOKAKTAEATPPRK 660  
 DB 601 GNRRTANVAHGFANLLTLDKKTLOEILVHVYPDSERIIMKKARVLLKOKAKTAEATPPRK 660  
 QY 661 DLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREAAOKKENSEGEGEGKENEJCL 720  
 DB 661 DLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREAAOKKENSEGEGEGKENEJCL 720  
 QY 721 KENEDKQKENEKDKENEDKDKGREPEEKPLDRPECTASPIAVEEPEHPSVRRVTLPRGTS 780  
 DB 721 KENEDKQKENEKDKENEDKDKGREPEEKPLDRPECTASPIAVEEPEHPSVRRVTLPRGTS 780  
 QY 781 ROSLIISMAPSAGGEEVLTIEVKEKAKQ 809  
 DB 781 ROSLIISMAPSAGGEEVLTIEVKEKAKQ 809  
 RESULT 2  
 ID CNGB3\_CANPA STANDARD; PRT; 782 AA.  
 AC Q8MJD7;  
 DT 29-MAR-2004 (Rel. 43, Created)  
 DT 29-MAR-2004 (Rel. 43, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Cyclic-nucleotide-gated cation channel beta 3 (CNG channel beta 3)  
 DE (Cyclic nucleotide gated channel beta 3) (Cone photoreceptor cGMP-  
 DE gated channel beta subunit) (Cyclic nucleotide-gated cation channel  
 DE modulatory subunit).  
 GN Name=CNGB3;  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;  
 OC Canis.

NCBI\_TaxID=9615;

[1]

NR NUCLEOTIDE SEQUENCE [MRNA], AND VARIANT CD ASN-262.  
PubMed=12140185; DOI=10.1093/hmg/11.16.1823;  
Sridjanin D.J., Lowe J.K., McElwee J.L., Milne B.S., Phippen T.M.,  
Sargan D.R., Aguirre G.D., Acland G.M., Ostrander E.A.;  
"the human CNGB3 mutations establish cone degeneration as orthologous to  
RT canine achromatopsia locus ACHM3.";  
Hum. Mol. Genet. 11:1823-1833 (2002).  
RL CC -1- FUNCTION: Visual signal transduction is mediated by a G-protein  
coupled cascade using cGMP as second messenger. This protein can  
be activated by cGMP which leads to an opening of the cation  
channel and thereby causing a depolarization of rod  
photoreceptors. Induced a flickering channel gating, weakened the  
outward rectification in the presence of extracellular calcium,  
increased sensitivity for L-cis diltiazem and enhanced the cAMP  
efficacy of the channel when coexpressed with CNGA3. Essential for  
the generation of light-evoked electrical responses in the red-,  
green- and blue sensitive cones (By similarity).  
RL CC -1- SUBUNIT: Heterooligomeric complex with CNGA3 (By similarity).  
RL CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
RL CC -1- DISEASE: Defects in CNGB3 are a cause of cone degeneration (cd).  
Cd is characterized by day-blindness and absence of retinal cone  
function. This autosomal recessive disorder occurs naturally in  
the Alaskan Malamute and German Shorthaired Pointer breeds.  
RL CC -1- SIMILARITY: Belongs to the cyclic nucleotide-gated cation channel  
(TC 1.A.1.5) family.  
RL CC -1- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.  
CC CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use as long as its content is in no way modified and this statement is not  
removed.

CC CC -----

DR ENBL; AF490511; AAM89224.1; -; mRNA.  
DR Ensembl; ENSCAFG0000008906; Canis familiaris.  
DR InterPro; IPR000595; CNMP\_bd.  
DR InterPro; IPR005821; Ion Trans.  
DR InterPro; IPR001622; K-channel\_pore.  
DR Pfam; PF00027; CNMP\_binding; 1.  
DR Pfam; PF00520; Ion\_trans; 1.  
DR SMART; SM00100; CNMP; 1.  
DR PROSITE; PS00888; CNMP\_BINDING\_1; 1.  
DR PROSITE; PS00889; CNMP\_BINDING\_2; 1.  
DR PROSITE; PS50042; CNMP\_BINDING\_3; 1.  
DR KW cGMP; cGMP-binding; Disease mutation; Ion transport; Ionic channel;  
KW Multigene family; Nucleotide-binding; Sensory transduction;  
KW Transmembrane; Transport; Vision.  
FT FT TOPO\_DOM 1 211 Cytoplasmic (Potential).  
FT FT TRANSHEM 212 232 H1 (Potential).  
FT FT TOPO\_DOM 233 245 Extracellular (Potential).  
FT FT TRANSHEM 246 266 H2 (Potential).  
FT FT TOPO\_DOM 267 297 Cytoplasmic (Potential).  
FT FT TRANSHEM 298 318 H3 (Potential).  
FT FT TOPO\_DOM 319 412 Extracellular (Potential).  
FT FT TRANSHEM 413 433 H4 (Potential).  
FT FT TOPO\_DOM 434 493 Cytoplasmic (Potential).  
FT FT TRANSHEM 494 514 H5 (Potential).  
FT FT TOPO\_DOM 515 572 Extracellular (Potential).  
FT FT TRANSHEM 573 593 H6 (Potential).  
FT FT TOPO\_DOM 594 782 Cytoplasmic (Potential).  
FT FT NP\_BIND 527 671 cGMP (By similarity).  
FT FT BINDING 587 587 cGMP (By similarity).  
FT FT BINDING 599 599 cGMP (By similarity).  
FT FT VARIANT 262 262 D -> N (in cd; in a German Shorthaired Pointer).  
SQ SEQUENCE 782 AA; 89424 MW; 7139EF3B86268A79 CRC64;

Query Match 75.1%; Score 3181.5; DB 1; Length 782;  
Best Local Similarity 76.2%; Pred. No. 3.3e-177;  
Matches 618; Conservative 57; Mismatches 105; Indels 31; Gaps 5;

NCBI\_TaxID=10090;  
 [1]  
 NUCLEOTIDE SEQUENCE, FUNCTION, SUBUNIT, AND TISSUE SPECIFICITY.  
 STRAIN=C57BL/6; TISSUE=Retina;  
 MEDLINE=20130348; PubMed=10662822;  
 Gerstner A., Zong X., Hofmann F., Biel M.;  
 "Molecular cloning and functional characterization of a new modulatory  
 cyclic nucleotide-gated channel subunit from mouse retina.";  
 J. Neurosci. 20:1324-1332(2000).  
 CC -!- FUNCTION: Visual signal transduction is mediated by a G-protein  
 coupled cascade using cGMP as second messenger. This protein can  
 be activated by cGMP which leads to an opening of the cation  
 channel and thereby causing a depolarization of rod  
 photoreceptors. Essential for the generation of light-evoked  
 electrical responses in the red-, green- and blue sensitive cones  
 (By similarity). Induced a flickering channel gating, weakened the  
 outward rectification in the presence of extracellular calcium.  
 CC increased sensitivity for L-cis diltiazem and enhanced the cAMP  
 efficacy of the channel when coexpressed with CNGB3.  
 CC -!- SUBUNIT: Heterooligomeric complex with CNGB3.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Small subset of retinal photoreceptor cells and  
 testis.  
 CC -!- SIMILARITY: Belongs to the cyclic nucleotide-gated cation channel  
 (TC 1.A.1.5) family.  
 CC -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use as long as its content is in no way modified and this statement is not  
 removed.  
 CC -----  
 DR EMBL; AJ243572; CAB71152.1; -; mRNA.  
 DR Ensembl; ENSMUSG0000056494; Mus musculus.  
 DR MGI; MGI:1353562; Cngb3.  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR GO; GO:0005886; C:plasma membrane; IC.  
 DR GO; GO:0030553; F:3',5'-cGMP binding; IC.  
 DR GO; GO:0005223; F:intracellular cGMP activated cation channel. . .; IPI.  
 DR GO; GO:0009187; P:cyclic nucleotide metabolism; IC.  
 DR InterPro; IPR000595; cGMP bd.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR001622; K-channel\_pore.  
 DR Pfam; PF00027; cNMP binding; 1.  
 DR Pfam; PF00520; Ion trans; 1.  
 DR SMART; SM00100; cNMP; 1.  
 DR PROSITE; PS00888; cNMP\_BINDING\_1; 1.  
 DR PROSITE; PS00889; cNMP\_BINDING\_2; 1.  
 DR PROSITE; PS00942; cNMP\_BINDING\_3; 1.  
 KW cGMP; cGMP-binding; Glycoprotein; Ion transport; Ionic channel;  
 KW Multigene family; Nucleotide-binding; Sensory transduction;  
 KW Transmembrane; Transport; Vision.  
 FT TOPO\_DOM 1 209 Cytoplasmic (Potential).  
 FT TRANSMEM 210 230 H1 (Potential).  
 FT TOPO\_DOM 231 242 Extracellular (Potential).  
 FT TRANSMEM 243 263 H2 (Potential).  
 FT TOPO\_DOM 264 294 Cytoplasmic (Potential).  
 FT TRANSMEM 295 315 H3 (Potential).  
 FT TOPO\_DOM 316 351 Extracellular (Potential).  
 FT TRANSMEM 352 372 H4 (Potential).  
 FT TOPO\_DOM 373 409 Cytoplasmic (Potential).  
 FT TRANSMEM 410 430 H5 (Potential).  
 FT TOPO\_DOM 431 558 Extracellular (Potential).  
 FT TRANSMEM 559 589 H6 (Potential).  
 FT TOPO\_DOM 590 694 Cytoplasmic (Potential).  
 FT NP\_BIND 524 668 cGMP (By similarity).  
 FT BINDING 584 594 cGMP (By similarity).  
 FT BINDING 596 596 cGMP (By similarity).  
 FT CARBOHYD 507 507 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 694 AA; 79722 MW; 0B9F9CF3B180DA82 CRC64;

Query Match

Best Local Similarity 68.7%; Pred. No. 2.3e-138;  
 Matches 485; Conservative 92; Mismatches 113; Indels 16; Gaps 7;  
 QY 1 MFKSLT-KVNVKPKPGENNEEQSRREEGSHP--SNOSQOTTAQEEKNKGEKSLKTKS 57  
 DB 1 MUKSLTVFKNVKNPM-----EGRMKKLCPNLSSLSQPTIAOGDNQSEKPLRSR- 50  
 QY 58 TPTVTEEPHTNTQDKLSKKNSSGDLTTPNDPQNAAEPTGTVPPEQKEMDPKGEKSPNQK 117  
 DB 51 TPTTEKSHSK-EDNSTGENSLRDTTPNDPECEAELETRTMAEMEKTRTKGERPVSFKTK 109  
 QY 118 PPAAPVINEYADAQHNLVKRMQRTALYKKKLVEGD-LSSPEASPTAKPTAVIPVKIS 176  
 DB 110 VLETSIINEYTDALHNLVERMRRTALYKKTLTTEENPPEVEASSOTAMSTN!SPKQFN 69  
 QY 177 DDKPTHEHYRLWPKVKMPLTEYLKRLKLPNSIDSYTDRLVLLWLLVLTAYNNCWFI 246  
 DB 170 NSKLKEH-QDTFSPAPQRPVKEHLRRMILPRSIDSYTDRVYLLWLLVLTAYNNCWIL 228  
 QY 237 PLRLVFPYQTADNIHYLIADIICDIYLYDMLFIQPRLQFVRRGGDIIVDSNELKHYRT 296  
 DB 229 PVRLVFPQTPDNKNYWIITDVCDIYLDLILQPRLOQFVRGGEIIVDSNELKRYRS 288  
 QY 297 STKFQDVASIIIPDICYLFFGFNPMFRANRMLKYTSFFEFNHHLESIMDKAYIYRVIRT 356  
 DB 289 STKFRMDVASLLPPELVYIFFGVNPIFRANRILKYTSFFEFNHHLESIMDKAYVYRVIRT 348  
 QY 357 TGYLLFIHLINACVYVWASNVGEGIGTRVWVDGEGEYLRVYVAVRVLITIGGLPEPOT 416  
 DB 349 TGYLLFIHLINACVYVWASDYEGIGTRVWVNGEGNLYRCYVAVRVLITIGGLPEPOT 408  
 QY 417 LFEIVFQLNFFSGVVFVSSLLIGQMRDVI GAATANQNYFRACMDDTIAYMNNYSIPKLVO 476  
 DB 409 SFEIVFQLNFFSGVVFVSSLLIGQMRDVI GAATANQNYFQACMDHI IAYMNNYSIPQSVQ 468  
 QY 477 KVRTWYETWDSQRMDESLLKTLPTTVQALAIQVNFVSIISKVDLIFKGCOTOMIDYM 536  
 DB 469 YRVRTWLEYTWNQSRILDESNNLENLPTAMOLSIALDINFSDIKVELFKGCDTOMIDL 528  
 QY 537 LLRLKSLVLYPGDFVCKKGEIGKEMWIIKHGEVQVLGGPDGTPKVLVTLKAGSVFGEISLL 596  
 DB 529 LLRLKSTIYLPQDFVCKKGEIGKEMWIIKHGEVQVLGGPDGGAQVLTUAKAGSVFGEISLL 588  
 QY 597 AAGGNNRRTANVAHGFANLLTLDKKTQLQELIVHYPSERILMKKARVLLKOKAKTAEAT 656  
 DB 589 AKGGNNRRTADVVAHGFANLLTLDKKTQLQELIHYPTSKLLMKKAKILLKOKGKTTOAI 648  
 QY 657 PPRKDLALLFPKKEPTPKLFTLLCGTCKASLARLLKLRQAAOK 702  
 DB 649 PARPGAPAFPPKKEPTPRMLKVLGNTGKVDLGRLLKGRKRTTQK 694  
 RESULT 4  
 Q80XL8\_MOUSE PRELIMINARY; PRT; 866 AA.  
 ID Q80XL8\_MOUSE PRELIMINARY; PRT; 866 AA.  
 AC Q80XL8;  
 DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DE Cngb1b protein (Fragment).  
 GN Name=Cngb1b;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Eye;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Query Match	46.6%	Score 1971.5	DB 2	Length 866
Best Local Similarity	47.5%	Pred. No. 1.7e-106		
Matches	406	Conservative 107	Mismatches 247	Indels 95
Gaps	14			
Qy	21	EOSRRNEGSHPSNQSOOTTAQENKEEKSLSKTSTPTVSEBPHNTIOBKLSKKNSSG	80	
Db	38	EEKEEKEEKEEKEEKEEKEEVEKEEKEEATNSTVPATKEPELOVED---TDADSG	94	
Qy	81	DL-----TTNPPDPQNAABPTQTVPEQKEMDPGKGGPN-----SPQKPPAA	121	
Db	95	PLIPEETLPPPERPPSPVKSDTLTPGAAAAAGHRKKLP SQDDEAEELKALSPAESPVA	154	
Qy	122	---PVINEYDAQ-----LHNI VKRMQROTALYKKLVGGDLSSPAS	161	
Db	155	WSDPTTPOAQDQDRAASTASQNSAI INDRLOELVKQPKETKEVKKEKLIIDPDVDEES	214	
Qy	162	PQAKPTAVPVVKGSDDKPT-----EHYYRLLFVKQKMPLETKRIKLPSNIDSYT	214	
Db	215	PKPS-PAKKAPEPDAQPAEAEVAEESHYCDMLCKFKRRPLKMY----RFPQSIDPLT	269	
Qy	215	DRLYLWLLLVTLAVNNNCWPIPLRLVPYQTADNIHWMLTADIICDIILYDMLPIQPR	274	
Db	270	NLMYILWLFVVLAWNNCWLIPLVRWAPPYQRAONIHFWLMDLYLDCDFIYLLDITVQMR	329	
Qy	275	LQFVRGGDIIIVDSNELRKHYRTSTKFDQDVASIIIPFDICYLFFGFNPMFRANRMLKYTSF	334	
Db	330	LQFVKGGDIIIDKKEMRNLYKSRRFKMDLCLPLDLFLYKLGINPLRLPRCLKYNAF	389	
Qy	335	FEFNHLESIMDKAVIYRVITTTGYLLFILMINACVYVWASNYEGIGTRWYDGEQNEY	394	
Db	390	FEFNRLBAIISKAVYRVIRTTAYLLYSLHNSCLTYWASAFQIGSTHWYDVGNSY	449	

```

RESULT 5
ID ID035788 RAT PRELIMINARY; PRT; 1339 AA.
AC O35788;
DT DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Cyclic nucleotide-gated channel beta subunit.
GN Name=Cngbl; Synonyms=CNMG4.1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX [1]
RN NUCLEOTIDE SEQUENCE. RP
RC STRAIN=Sprague-Dawley;
RA Sautter A., Biel M., Hofmann F.;
RT "Molecular cloning of cyclic nucleotide-gated cation channel subunits
RT from pineal gland.";
RL EMBL; AJ000496; CAA04133.1; -; mRNA.
DR HSSP; O88703; I043
DR Ensembl; ENSRNOG0000013544; Rattus norvegicus.
DR RGD; 621809; Cngbl.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR000595; cNMP_binding.
DR InterPro; IPR003821; ion_trans.
DR Pfam; PF00027; cNMP_binding; 1.
DR Pfam; PF00520; Ion_trans; 1.
DR SMART; SM00100; cNMP; 1.
DR PROSITE; PS00888; CNMP_BINDING_1; UNKNOWN_1.
DR PROSITE; PS00889; CNMP_BINDING_2; 1.
DR PROSITE; PS50042; CNMP_BINDING_3; 1.
DR SQ SEQUENCE 1339 AA; 151046 MW; FFAD64A6A81AB49A CRC64;

```

Query Match 46.3%; Score 1959; DB 2; Length 1339;

Best Local Similarity 46.3%; Pred. No. 1.6e-105; Matches 400; Conservative 121; Mismatches 240; Indels 102; Gaps 15;	
QY 11 VKPIGNNENEGSSRRNEGSHPSNQSQTTAQEENKGEKSLKTKSTPTVTEEPHTNIQ 70	OS
Db 512 IQELPEEEKEEKEEKEE---BEKEEKEEKEEKEEKEEATSTVPATKEHPELQVE 568	OC
QY 71 DKLSKNSSGDIITNPPDQN-----AAEPTGTVP-----BQKEMDPK 108	OC
Db 569 DTDAAAGPLIPETITPPPPVPSPAKSDTLAVPSAATHRKXLPQDDAEELKALSPAE 628	OC
QY 109 -----EGNSPQ-----NKPAPVINEVADAOLHNLVKMRORTALYKKLVE 152	OC
Db 629 SPVVAWSDDTTTQADGEDRAASTASQNSAIIND-----RLQELVKMKFKERTKVKELJD 684	OC
QY 153 GDLS-----SPEASPTAKTAVPPVK--ESDDKPTHEYVRLWLFVKKMPLEYLKRKL 206	OC
Db 685 PVTSDSESPKSPAKAPDSAPAKPAEAAEAEHCDMLCKCKFKRPPWKY-----QF 740	OC
QY 207 PMSIDSYDRLYLLMLLVTLAYNNWCWFIPLRLVFPYQTADNIHYWLIADIICDIYLY 266	OC
Db 741 PQSIDPLTNLMYLMVLFVFLAWNNCWLIPVRWAPPYQADNIHLWLLMDVLCDFIYL 800	OC
QY 267 DMLFTQPRLOFVRGGDIIVDSNELKHYRTSTKFLQDVASIIPDFICYLFFGPNMFRAN 326	OC
Db 801 DITVFQMRLOFVKGGDIITDKKEMRNLYLKSQRFKMDLCLLPDLYLKLGVNPLLRLP 860	OC
QY 327 RMLKTYSPFENHLESIMDKAYIVRVRTTGVLLFILHINACVYVWASNYEGIGTRWV 386	OC
Db 861 RCLKYWAFEFNRLAELLSKAYVYVIRTTAYLLYSLHNSCLYVWASAFQIGSTHWV 920	OC
QY 387 YDGEYELRCYVWAVRTLITIGLPEPQTLFEIVFQLNFFSGVVFVSSLIGQMRDVTG 446	OC
Db 921 YDGVNGVYRCYVWAVKTLITIGLPPDQTLFEIVFQLNFFSGVVFVSSLIGQMRDVTG 980	OC
QY 447 AATANQYFRACWDOTIAMNNYSIPKLVKRVRTWYETWDSQRMDESLLKTLPTTV 506	OC
Db 981 AATAGTYRSCWDSVTVMNFYKIPRSVQNRVKTWYETWDSQRMDESLLKTLPTTV 1040	OC
QY 507 QALADVNFSIISKVDLPKGGDTOMIYDMLRLKSLVLPDGFVCKKGEIGKEMVYIKH 566	OC
Db 1041 RLDAIDVNNIVSKVAFQGGDRQMIYDMLRLKSLVLPDGFVCKKGEIGKEMVYIKH 1100	OC
QY 567 GEVQVLGGPDGKVLVTLKAGSVFGEISLLAAGGNNRRTANVVAHFANLLTLDKKTLOE 626	OC
Db 1101 GQVQVLGGPDGKVLVTLKAGSVFGEISLLAAGGNNRRTANVVAHFANLLTLDKKTLOE 1160	OC
QY 627 ILVHYVDSERILMKARVLLKOKATAEATPPKOLALLFPKKEETPKLFTLLGGTGKA 686	OC
Db 1161 ILVHYVDSERILMKARVLLKOKATAEATPPKOLALLFPKKEETPKLFTLLGGTGKA 1215	OC
QY 687 S-----LARLLKKEQAA-----QKENSEGEEBEGKENEKQENEDK 726	OC
Db 1216 GPRGAKGKLAHLRLKELALEAARQOQLLEQAKSQEAGGEGSGATQPPAPQEPS 1275	OC
QY 727 QKENEDKGENEDKQGRPEBKPLDRPECTSPAVEBEPHSVRTVLPRTGSRQSLII 786	OC
Db 1276 -----EPKEPP--EPPAPSSPPPSAKPEG--STEEAAGPPEPSVRI 1313	OC
QY 787 SNAPSAGGEEVLTIEVKEKAKQ 809	OC
Db 1314 RVSPGDPGEQTLVSEMLEEKKE 1336	OC
RESULT 6	
O43636 HUMAN	
ID O43636 HUMAN PRELIMINARY; PRT: 1245 AA.	
AC O43636;	
DT 01-JUN-1998 (Tremblrel. 06, Created)	
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)	
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)	
DE Rod photoreceptor CNG-channel beta subunit.	
GN Name=RCNC2;	

OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC	Homo.
OX	NCBI_TaxID=9606;
RP	NUCLEOTIDE SEQUENCE.
RA	Grunwald M.E., Yu W.P., Yu H.H., Yau K.W.;
RT	"Identification of a domain on the beta subunit of the rod cGMP-gated
RT	cation channel that mediates inhibition by calcium-calmodulin.";
RL	J. Biol. Chem. 0:0-0(1998).
RN	NUCLEOTIDE SEQUENCE.
RP	MEDLINE=96338110; PubMed=8766832; DOI=10.1016/0014-5793(96)00588-1;
RX	Ardeli M.D., Aragon I., Oliveira L., Porche G.E., Burke E.,
RA	Pittler S.J.;
RT	"The beta subunit of human rod photoreceptor cGMP-gated cation channel
RT	is generated from a complex transcription unit.";
RL	FEBS Lett. 389:213-218(1996).
DR	EMBL; AF042498; AAC04830.1; -; mRNA.
DR	PIR; S32538; S32538.
DR	PIR; S69275; S69275.
DR	HSSP; O88703; I043.
DR	Ensembl; ENSG0000070729; Homo sapiens.
DR	GO; GO:0015276; F:ligand-gated ion channel activity; TAS.
DR	GO; GO:0006810; P:transport; TAS.
DR	GO; GO:0007601; P:visual perception; TAS.
DR	InterPro; IPR000595; cNMP binding.
DR	InterPro; IPR005821; Ion trans.
DR	InterPro; IPR001622; K-channel_pore.
DR	Pfam; PF00027; cNMP binding; 1.
DR	Pfam; PF00520; Ion trans; 1.
DR	SMART; SM00100; cNMP; 1.
DR	PROSITE; PS00889; cNMP_BINDING_1.
DR	PROSITE; PS00889; cNMP_BINDING_2; 1.
DR	PROSITE; PS00042; cNMP_BINDING_3; 1.
DR	Receptor.
QY	SEQUENCE 1245 AA; 139160 MW; 40C4860BFCF86126 CRC64;
Query Match 46.1%; Score 1952; DB 2; Length 1245;	
Best Local Similarity 46.4%; Pred. No. 3.7e-105;	
Matches 405; Conservative 121; Mismatches 237; Indels 110; Gaps 18;	
QY 14 IGENNENQSSRRNEEGSHPSNQSQTTAQE--ENKGE--EKSUKTKSTPTVTEEPHTN 68	Db
403 VGEEAKAEAEKAE 462	
69 IQDK-----LSKNSSGDLTTNPDQNAAEPTCTVPEQKEMDPCKEGPN----- 112	
463 VEDTDADSCPLMAEENPPS--TVLPPSPAKSDTLIVPSSASGTHRKKLPSEDDAEELK 520	
113 --SPONKP-----PAAP-----VINEYADAOLHNLVKMRORTALY 146	
521 ALSPAESPVAWSDDTTTQDGDQDRAASTASTNSAIIND RLQELVKLFKERTKVK 576	
147 RKKLVEGDL-----SPEASPTAKTAVPPVKESDDKPT--EHYVRLWLFVKKMPLEY 200	
577 KEKLDIPDVTSDSESPKSPAKAPAPADTKPAEAEPEVEEHCYCDMLCKCKFKRPPWKY 636	
201 LKRIKLPNSIDSYDRLYLLMLLVTLAYNNWCWFIPLRLVFPYQTADNIHYWLIADIIC 260	
637 ---QFQSIDPLTNLMYLMVLFVFLAWNNCWLIPVRWAPPYQADNIHLWLLMDVLC 692	
261 DIIVLYDMLFTQPRLOFVRGGDIIVDSNELKHYRTSTKFLQDVASIIPDFICYLFFGPN 320	
693 DLIVFLDITVQTRLOFVRGGDIITDKKEMRNLYLKSRRFKMDLCLLPDLYLKGVN 752	
321 PMFRANRLKYTSFFENHLESIMDKAYIVRVRTTGVLLFILHINACVYVWASNYEGI 380	
753 PLLRLPRCLKYMAFPFENSRLESILSKAYVYVIRTTAYLLYSLHNSCLYVWASAYQGL 812	
381 GTTRWYDGEYELRCYVWAVRTLITIGLPEPQTLFEIVFQLNFFSGVVFVSSLIGC 440	

```

Db      813  GSTHWYDGVGNSYIRCYFAVKTLITIGGLPDPKTLFRIVFOLLNFTGVPFASVWIGQ 872
Qy      441  MRDVIQAATANYFRACMDDTIAYNNYSIPKLVQKRVRTWYTYWDSQRMLEDSDLLK 500
Db      873  MRDVGAAATAGQTYRSCMDSTVKYNNFYKIPKSVQNRVKTWYTYWHSQGMLEDELMV 932
Qy      501  TLPITVQLALADINFSIISKVDLPKGCOTOMYDMLRLKSLVLYLPDGFVCKKGIGKE 560
Db      933  QLPDKRLDLADINVTNIVSKVALFQGCQRMIFDMLKRLRSVVYLPFDVYCKKGIGRE 992
Qy      561  MYIIRKEGVQVLGGPDTKVLVTLTKAGSVFGEISLLAAGGGRNRTANVAHGFANILLTD 620
Db      993  MYIIQGVQVLGGPDKSVLTVLKAGSVFGEISLLAAGGGRNRTANVAHGFANILLTD 1052
Qy      621  KKTQLBIHVHPDSEIRILMKKARVLLKQAKATAETPPRKDLALLPPPKETPKLFTLL 680
Db      1053  KDLNEILVHPSPQKLLRKARRMLRSNNK-----PKEKSVLILPPRAGTPKLFNAAL 1107
Qy      681  GGTGKAS-----LARIILKREQAOKKENSEGEGEGEKENEDKQK 728
Db      1108  AMTGRMGKGAGKGLAHLRLKELAALEAAKQQL-----VEQAKSSQDVKGEGSAAP 1164
Qy      729  ENEDKGE-NEDKDKGREPEK-----LDRPECTASPIAVEEPHPSVRRTVLP 776
Db      1165  DQHTHPKEATDPPARTPEPPGSPSSPPASLGRPEGEERGPA-SPEHSVR----- 1218
Qy      777  RGTROSRLIISMAFSGEGSEVLTIEVKEKAKQ 809
Db      1219  -----ICMSGPGEPGQILSVKMPERE 1242

RESULT 7
CNGB1 HUMAN
ID CNGB1 HUMAN STANDARD; PRT; 909 AA.
AC Q14028; Q14029;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
DE Cyclic-nucleotide-gated cation channel 4 (CNG channel 4) (CNG-4)
DE (CNG4) (Cyclic nucleotide-gated cation channel modulatory subunit).
GN Name=CNGB1; Synonyms=CNGC4;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORMS RCNC2A AND RCNC2B).
RC TISSUE=Retina;
RX MEDLINE=93226050; PubMed=7682292; DOI=10.1038/362764a0;
RA Chen T.Y., Peng Y.-W., Dhallan R.S., Ahamed B., Reed R.R., Yau K.-W.;
RT "A new subunit of the cyclic nucleotide-gated cation channel in
RL retinal rods";
RL Nature 362:764-767(1993).
CC -!- SUBUNIT: Heterooligomeric complex with CNG1.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=RCNC2B;
CC IsoId=Q14028-1; Sequence=Displayed;
CC Name=RCNC2A;
CC IsoId=Q14028-2; Sequence=VSP_001110;
CC -!- SIMILARITY: Belongs to the cyclic nucleotide-gated cation channel
CC (TC 1.A.1.5) family.
CC -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; L15296; AAA65620.1; -; Genomic_DNA.

```

```

DR      14  IGGENNEQSSRRNEEGSHPSNOSQQTAAQE--ENKGE---EKSLTKTSPVTSEEPHTN 68
DR      67  VGEAKKEAEKAEAEVAAEEKEPODWAETKEEPEAEAEAAASSGVPAATKQHPVQ 126
Qy      69  IQDK-----LSKKNSSGDLTTNPQNAAEPTGTVPCKMDPGKEGPN----- 112
Db      127  VEDTDADSCPLMAEENPPS--TVLPSPSPAKSDTLIVPSSASCTHRKKLPSEDEAEELK 184
Qy      113  --SPQNKP-----PAAP-----VINEYADAQLHNLVKKRMRORTALY 146
Db      185  ALSPAESPVVAVSDPTTPKTDGQDRAASTASTNSAIIND----LQELVKLFKERTKV 240
Qy      147  KKLVGEGDLS-----SPEASQTAKTATVPVKESDDKPT--EHYRLLWFKVKMKPLTEY 200
Db      241  KEKLIIDPDVTSDEESPSPAKKAPADPTKPAEAPVEEHEYCDMLCKCFKHPWKKY 300
Qy      201  LKRIKLPNISIDSYTRLLYLWLLVLTALYNWNCWFPLRLVFPVQYADNLIHYMLADIIC 260
Db      301  ----QFPQSIDPLTNLMYLVWLFVVMWNNWNCWLPVRWAFPYQTPDNIHHMLMDYLC 356
Qy      261  DIILYDMLFIQRLQFVRGGDIIVDSNELRKHYRTSTKFKQDVASIIIPEDICYLFFGPN 320
Db      357  DLIVFLDITVFQRLQFVRGGDIITDKDMRNLYKSRFRKMDLLSLPLDLFLYLVKGVN 416
Qy      321  PMFRANRLKYTSFFPENHLESIMDKAYIRVIRTYGTYLLFILHINACVYVWASNYEGI 380
Db      417  PLRLPRCLLYMAFFEFNSRLESILSKAYVYRIVRTAYLLYLSLHLSCLYVWASAYOGL 476
Qy      381  GTTRWYVYDGRGNEVLCYVWAVRTLIITIGGLPDPKTLFRIVFOLLNFTGVPFASVWIGQ 440
Db      477  GSTHWYDGVGNSYIRCYFAVKTLITIGGLPDPKTLFRIVFOLLNFTGVPFASVWIGQ 536

Query Match      46.0%; Score 1949; DB 1; Length 909;
Best Local Similarity 46.2%; Pred. No. 3.6e-105;
Matches 403; Conservative 122; Mismatches 238; Indels 110; Gaps 18;

Qy      14  IGGENNEQSSRRNEEGSHPSNOSQQTAAQE--ENKGE---EKSLTKTSPVTSEEPHTN 68
Db      67  VGEAKKEAEKAEAEVAAEEKEPODWAETKEEPEAEAEAAASSGVPAATKQHPVQ 126
Qy      69  IQDK-----LSKKNSSGDLTTNPQNAAEPTGTVPCKMDPGKEGPN----- 112
Db      127  VEDTDADSCPLMAEENPPS--TVLPSPSPAKSDTLIVPSSASCTHRKKLPSEDEAEELK 184
Qy      113  --SPQNKP-----PAAP-----VINEYADAQLHNLVKKRMRORTALY 146
Db      185  ALSPAESPVVAVSDPTTPKTDGQDRAASTASTNSAIIND----LQELVKLFKERTKV 240
Qy      147  KKLVGEGDLS-----SPEASQTAKTATVPVKESDDKPT--EHYRLLWFKVKMKPLTEY 200
Db      241  KEKLIIDPDVTSDEESPSPAKKAPADPTKPAEAPVEEHEYCDMLCKCFKHPWKKY 300
Qy      201  LKRIKLPNISIDSYTRLLYLWLLVLTALYNWNCWFPLRLVFPVQYADNLIHYMLADIIC 260
Db      301  ----QFPQSIDPLTNLMYLVWLFVVMWNNWNCWLPVRWAFPYQTPDNIHHMLMDYLC 356
Qy      261  DIILYDMLFIQRLQFVRGGDIIVDSNELRKHYRTSTKFKQDVASIIIPEDICYLFFGPN 320
Db      357  DLIVFLDITVFQRLQFVRGGDIITDKDMRNLYKSRFRKMDLLSLPLDLFLYLVKGVN 416
Qy      321  PMFRANRLKYTSFFPENHLESIMDKAYIRVIRTYGTYLLFILHINACVYVWASNYEGI 380
Db      417  PLRLPRCLLYMAFFEFNSRLESILSKAYVYRIVRTAYLLYLSLHLSCLYVWASAYOGL 476
Qy      381  GTTRWYVYDGRGNEVLCYVWAVRTLIITIGGLPDPKTLFRIVFOLLNFTGVPFASVWIGQ 440
Db      477  GSTHWYDGVGNSYIRCYFAVKTLITIGGLPDPKTLFRIVFOLLNFTGVPFASVWIGQ 536

EMBL; L15297; AAA65619.1; -; Genomic_DNA.
Ensembl; ENSG00000070729; Homo sapiens.
HGNC; HGNC:2151; CNGB1.
MIM; 600724; -
GO; GO:0017071; C:intracellular cyclic nucleotide activated c.; -; NAS.
GO; GO:0005222; F:intracellular cAMP activated cation channel.; -; NAS.
InterPro; IPR0050595; C:NM b.
InterPro; IPR005821; I:NM trans.
InterPro; IPR001622; K:channel_binding; 1.
Pfam; PF00520; I:NM trans; 1.
PROSITE; PS00888; C:NM_BINDING_1; 1.
PROSITE; PS00889; C:NM_BINDING_2; 1.
PROSITE; PS00442; C:NM_BINDING_3; 1.
KW Alternative splicing; cAMP; cAMP-binding; Ion transport;
KW Ionic channel; Multigene family; Nucleotide-binding; Transmembrane;
KW Transport.
FT TOPO_DOM 1 314 Cytoplasmic (Potential).
FT TRANSMEM 315 333 H1 (Potential).
FT TOPO_DOM 334 347 Extracellular (Potential).
FT TRANSMEM 348 366 H2 (Potential).
FT TOPO_DOM 367 391 Cytoplasmic (Potential).
FT TRANSMEM 392 411 H3 (Potential).
FT TOPO_DOM 412 448 Extracellular (Potential).
FT TRANSMEM 449 471 H4 (Potential).
FT TOPO_DOM 472 515 Cytoplasmic (Potential).
FT TRANSMEM 516 535 H5 (Potential).
FT TOPO_DOM 536 619 Extracellular (Potential).
FT TRANSMEM 620 640 H6 (Potential).
FT TOPO_DOM 641 909 Cytoplasmic (Potential).
FT NP_BIND 628 767 cAMP (By similarity).
FT COMBIAS 9 29 Poly-Glu.
FT BINDING 688 688 cAMP (Potential).
FT BINDING 700 700 Missing (in isoform RCNC2A).
FT VARSPLIC 1 286 /FTId=VSP_001110.
SQ SEQUENCE 909 AA; 102285 MW; DC0E75433686EDDD CRC64;

```



QY	118	PPAA	AVINEYADAO	QLHNLV	KEMRORTAL	YKKKL	VEGDL	----	SPAS	POTAKPT	AVPPV	17																																											
Db	173	SONS	AIID	----	RLQEL	VKRFK	ERTK	VEK	KLID	PDVT	SDESP	KPKSPAK	ADPSAPAO	228																																									
QY	174	K	----	ESDD	KPT	EHYRL	LMFKV	KQKPL	TEYLK	RKL	PNSID	SYTD	RLYL	LMWLL	LVTL	AYNN	231																																						
Db	229	KPAE	AAAAA	EEHY	CDML	CK	KFR	PKWY	----	QFQ	SID	PL	NLMY	IL	ML	FFVL	AWNN	284																																					
QY	232	NCWF	IPLRL	VLP	QYAD	NH	WL	ADI	IC	DI	YL	YD	ML	F	Q	PL	RQ	LV	RG	DI	IV	DS	NELR	291																															
Db	285	NCWL	IP	VRA	WAF	PYQ	RAD	NH	LL	W	LD	Y	LC	F	I	Y	L	D	I	T	F	Q	M	R	L	Q	V	K	G	D	I	I	T	D	K	E	M	R	344																
QY	292	KHYR	ST	F	E	Q	L	D	V	A	S	I	I	P	D	I	C	Y	L	F	F	G	N	P	M	F	R	A	N	M	L	K	Y	T	S	P	E	F	F	N	H	L	E	S	I	M	D	K	A	Y	I	Y	351		
Db	345	NNYL	SQ	R	F	Q	M	D	L	C	L	P	E	L	F	Y	L	K	G	V	N	P	L	R	P	C	K	Y	M	A	F	E	F	E	N	R	L	E	A	L	S	K	A	Y	V	404									
QY	352	RVIR	T	G	Y	L	L	F	I	L	H	N	A	C	V	Y	W	A	S	N	E	G	I	T	T	R	W	V	D	E	G	N	E	Y	L	R	C	Y	Y	M	A	V	R	L	I	T	I	G	L	411					
Db	405	RVIR	T	A	V	L	L	S	L	N	S	C	L	Y	W	A	S	A	F	O	G	I	G	S	T	H	W	V	D	G	N	S	Y	I	R	C	Y	Y	M	A	V	R	L	I	T	I	G	L	464						
QY	412	PEP	Q	T	L	FE	I	V	F	O	L	L	N	F	F	S	G	V	S	L	I	G	Q	M	R	D	V	I	G	A	T	A	N	O	N	V	P	R	A	C	M	D	T	I	A	M	N	Y	S	I	471				
Db	465	PD	Q	T	L	FE	I	V	F	O	L	L	N	F	T	G	V	F	A	S	V	M	I	G	Q	M	R	D	V	I	G	A	T	A	N	O	N	V	P	R	A	C	M	D	T	I	A	M	N	Y	S	I	524		
QY	472	PKL	VQ	K	R	V	R	T	W	E	Y	T	W	S	O	R	M	L	D	S	L	L	K	T	L	P	T	T	V	O	L	A	I	D	V	N	F	S	I	S	K	V	D	L	F	K	G	C	D	T	531				
Db	525	PR	S	V	Q	N	R	V	T	W	E	Y	T	W	S	O	G	M	L	D	S	E	L	M	V	O	L	P	D	K	M	R	L	D	L	A	I	D	V	N	Y	I	V	S	K	V	A	L	F	O	G	C	D	R	584
QY	532	MYD	M	L	L	R	L	K	S	V	L	P	G	F	V	C	K	E	I	G	K	E	M	Y	I	I	K	H	G	E	V	O	V	I	L	G	P	D	T	K	V	L	T	L	K	A	G	S	V	F	591				
Db	585	MIF	D	M	L	K	R	S	V	V	L	P	N	D	V	C	K	E	I	G	R	E	M	Y	I	I	O	A	G	V	O	V	I	L	G	P	D	T	K	V	L	T	L	K	A	G	S	V	F	644					
QY	592	EIS	L	A	A	G	G	N	R	T	A	N	V	A	H	G	F	A	N	L	L	T	D	K	K	T	L	O	E	I	L	V	H	P	D	S	E	R	I	L	M	K	K	A	R	V	L	L	K	O	K	651			
Db	645	EIS	L	A	A	G	G	N	R	T	A	N	V	A	H	G	F	T	N	L	F	I	L	D	K	K	L	E																											

DR	GO:	GO:0016020; C-membrane; IEA.
DR	GO:	GO:0005216; F-ion channel activity; IEA.
DR	GO:	GO:0006811; P-ion transport; IEA.
DR	InterPro:	IPR000595; cNMP_binding.
DR	InterPro:	IPR005821; Ion_trans.
DR	Pfam:	PF00027; cNMP_binding; 1.
DR	Pfam:	PF00520; Ion_trans; 1.
DR	SMART:	SM00100; cNMP; 1.
DR	PROSITE:	PS00886; CNMP_BINDING_1; UNKNOWN_1.
DR	PROSITE:	PS00889; CNMP_BINDING_2; 1.
DR	PROSITE:	PS00442; CNMP_BINDING_3; 1.
SQ	SEQUENCE	938 AA; 104518 MW; DIF4PBCF18E53EC4 CRC64;

  

Query Match		45.3%; Score 1920; DB 2; Length 938;
Best Local Similarity		45.9%; Pred. No. 1.9e-103;
Matches 401; Conservative 118; Mismatches 242; Indels 112; Gaps 17;		

  

Qy	15	GNNENFGSSRRNEGSHPSNQSQTTAAENKGEBSLKTKSTPTVTSBETHIQDK--	72
Dd	76	GAQAQGEVGGAQEODGVGGA-QDQSTSHQELQ--EALADSGVPATEHPELQVEDADA	132
Qy	73	-----LSKNSSGDLTTPNPQNAAEPGTGPV-----EQEMDPGK	108
Dd	133	DSRPFLIAENPPSPVLPLSP-AKSDTLAVPGSATGSLRKRLSPQDDAEELKWLPAA	190
Qy	109	-----EGPNSPONKPPA-----APVINEYADAQLNLVKRMQRQTALYKKKLAVE	152
Dd	191	SPVVAWSDPSTSPQGTDDQDRATSTASQNSAIIND----RLQELVKLFKERTEKVKEKLID	246
Qy	153	GDLS-----SPASQTAKTAVPVVK--ESDDKPTHEYHYLLPKVKOMPUTEYLKRILK	206
Dd	247	PDVTSDEESPSPAKKAPEAPEVKPAEAQGVBEEHYCEMLCKCFKRRPWKKY----	302
Qy	207	PNSTDSTVDRLYLWLLLVTLAYANNWCWFILRLVFPVQTDADNIHMYLIADIICDIIVLY	266
Dd	303	QSIDPTNLMIYIIWLFPVLAWNNWCMLIIPRVAFPYQTDPDNHLMLMDYLCDLIYLL	362
Qy	267	DMLPIQRLQVRGGDIIVDSNELRKHYRTSTKFQLDVASITPEDICYLFFGFNPMFRAN	326
Dd	363	DITVFQMLQFVRGGDIITDKERNNYVKSQRFKMOMCLLPLDLLYLFKEGVNPLLRLP	422
Qy	327	RMLKYTSFFFNHHLESIMDKAYIRVURTGTGYLLFILHNACVYVNASNTVEGITTRTW	386
Dd	423	RCLYMAFEFENNRESILSKAYYRVIRTAYILYLSHLNSCLYVNASAYEGLGSTHW	482
Qy	387	YDGSENEVLRVCYWAVRTLTIIGLPEPOTLFEIVFOLLNFPSGVFSSILGOMRDVIG	446
Dd	483	YDGVNSYIRCYWAVKXLTITIGLDPDLTLFEIVFQGLANTFTGVFAFSVMIGMRDVG	542
Qy	447	AATANQYFRACMDDTIAYNMNYGI PKLVQKRVRTWEYTWD SQRMLDDESLLTLP TTV	506
Dd	543	AATAGQTYRYSCMDBSTVKYMNFKI PRSVQNRVKTWETWH S QGM L D E S E L A W Q L P D K M	602
Qy	507	QLALAIDVNPISII SKVDFKGCDQM YDM LL R L K S V I L Y L P G D F V C K G E T G K M Y I I K H	566
Dd	603	R L D L A I D V N Y S I V S K V A L F O G C D R Q M I F D M L K R L S V Y L E N D V V C K G E I G R E M Y I I Q A	662
Qy	567	GEVOVLGPDGTKVLVTIKAGSVFGEISLLAAGGNRTANNVAHGFANILLTLOKKTLOE	626
Dd	663	GQVQVLGPDGKSVLVTIKAGSVFGEISLLAAGGNRTANNVAHGFNTLFILOKDOLNE	722
Qy	627	ILVHPDSEIRLMKARVLLKQAKTABEATPRPKDALLFPFKBETPKLFKTLILGGTGKA	686
Dd	723	ILVHPSPQKLLRKARMLRNKK-----PKESVLIILPPRACTPKLFXAALAAAGKM	776
Qy	687	SL-----ARLLKREQAOKKENSEGEBEGKENEKQENEDKQENEDBKQ	734
Dd	777	GAKGGRGRIALLRLARKELAALAAEARQQOL---LEQAKSSEDAAVGEE-----G	824
Qy	735	KENEDKDKGREPEEKPLDRPECTA-----SPATVEEBSPHSVRRTVLP	776
Dd	825	SASPEQPRPREPEAPEAPEPTAPEPLAPEAPEAPAPASSPPAPSOERPEGDKDAARP	884

```
QY 387 YDGEYNEYLRCYVWAVRTLTITIGLPEPOTLFEIVFOLLNFFSGVFVFSLSLQMRDVTG 446
Db 493 YDGVGNSYRCYVWAVRTLTITIGLPEPOTLFEIVFOLLNFFSGVFVFSLSLQMRDVTG 552
QY 447 AATANQYPRACNDTTIAYNNYSIPKLQKRVRTWYEVTSORSMLDESLLKTLPTTV 506
Db 553 AATAGQTYRSCMDSTVYKYNFYKIPRSVQNRVKTWYEVTSORSMLDESLLKTLPTTV 612
QY 507 QLALAIADVNFISIISKVDLFGKCDTOMIYDMLRLKSVLYLPGDFVCKGGEIGKEMIIKH 566
Db 613 RLDAIDVNYISVSKVALFGCDROMIFDMLKRLSVLYLPGDFVCKGGEIGKEMIIKH 672
QY 567 GEVQVLGGPDGKVLVTLKAGSVFGEISLLAAGGNNRTANVVAHFANLLTLDKKTLOE 626
Db 673 GQVQVLGGPDGKVLVTLKAGSVFGEISLLAAGGNNRTANVVAHFANLLTLDKKTLOE 732
QY 627 ILVHYPDSEIRILMKARVLLKQAKTAEATPPRKDLALLFPKKEETPKLFTLLGGTGKA 686
Db 733 ILVHYPDSEIRILMKARVLLKQAKTAEATPPRKDLALLFPKKEETPKLFTLLGGTGKA 786
QY 687 SL-----ARLLKLRQAAQKENSSEEGEENEDKOKENEDKOKENEDK 734
Db 787 GAKGGRGRLALLRLKELALEAAARQOOL-----LEQAKSDEDAVGE-----G 834
QY 735 KENEDKDKGREPEKPLDRPECTA-----SPIAVEEPPHSVRRTVLP 776
Db 835 SASPEQPPRPEPAPEAPEPTAPEPAPEAPEAPEAPEAPEAPEAPEAPEAPEAPEAPEAPE 894
QY 777 RGTSRQSLIISMAPSAGGGEVLTIEVKEKAKQ 809
Db 895 ---EEHPVRIHVTGLGDPDPSEQILLVEVPEKQEE 924

RESULT 12
O77660 BOVIN PRELIMINARY; PRT; 952 AA.
ID AC O77660 BOVIN PRELIMINARY; PRT; 952 AA.
AD O77660;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Cyclic nucleotide-gated channel beta subunit 1e.
GN Name=CNCbeta;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98345361; PubMed=9679145; DOI=10.1083/jcb.142.2.473;
RA Wiesner B., Weiner J., Middendorff R., Hagen V., Kaupp U.B.,
RA Weyand I.;
RT "Cyclic nucleotide-gated channels on the flagellum control Ca2+ entry
into sperm."
RL J. Cell Biol. 142:473-484(1998).
DR EMBL; AF074014; AAC26129.1; -; mRNA.
DR HSP; O88703; 1043.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR000595; cNMP binding.
DR InterPro; IPR005821; Ion trans.
DR Pfam; PF00027; cNMP binding; 1.
DR Pfam; PF00520; Ion_trans; 1.
DR SMART; SM00100; cNMP; 1.
DR PROSITE; PS00886; cNMP_BINDING_1; UNKNOWN_1.
DR PROSITE; PS00889; cNMP_BINDING_2; 1.
DR PROSITE; PS00442; cNMP_BINDING_3; 1.
DR PROSITE; PS00442; cNMP_BINDING_4; 1.
SQ SEQUENCE 952 AA; 105913 MW; D43F8B1D27F7F052 CRC64;

Query Match 45.3%; Score 1920; DB 2; Length 952;
Best Local Similarity 45.9%; Pred. No. 1.9e-103;
```

```
Matches 401; Conservative 118; Mismatches 242; Indels 112; Gaps 17;
QY 15 GENNESSRRNNEGSHPSNQSQTAAQENKNGEKSUKTSTPTVTSEPTNIQDK-- 72
Db 90 GAQAQGEVGAQEQDVGGA-QDQSTSHOEL--BEALADSSGVPATEHPELQVEDADA 146
QY 73 -----LSKNSSGDLTTNPDPQNAAEPTCTVP-----BOKEMDPCK 108
Db 147 DSRPLIAEENPPSPVOLPLSP--AKSDTLAVPGSATGSLRKRLPSODDDAEELKMLSPAA 204
QY 109 -----EGPNSPQNKPPA-----APVINEYADAQHLNLVKRMORTALYKKKLVE 152
Db 205 SPVWMSDPTSFGTDDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQ 260
QY 153 GDLS-----SPEASPTAKPTAVPPVK--ESDDKPTHEYRLLWFKVKMPLTEYLKRIKL 206
Db 261 PDVTSDEESPSPAKAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAPAP 316
QY 207 PHSIDSYDRLVLLMLLVTLAYNNCWFIPLRVFPYOTADNIHYWLIADIICDIILY 266
Db 317 POSIDPLTLMLYILMLFFVLAWNWCWLIPLRVAPPYOTPDNIHLWLMVLCDLIYLL 376
QY 267 DMLFTQPLQFVRGGDIIVDSNELKHYRTSTKFLQDVASIIIPFDICYLFFGFNPMFRAN 326
Db 377 DITVFQMRLOFVRGGDIITDKEMRNYYKVSORFKWDMCLCLPLDLLLYLFKGVNPLRLP 436
QY 327 RMLKYTSPFEPNHLIESIMDKAYIYRVIRTTGVLFIPLHINACVYVWASNYEGIGTRVW 386
Db 437 RCLKYMAPFEPNHLIESILSKAYVYRVIRTTAYLLYSLHNSCLYVWASNYEGIGSTHW 496
QY 387 YDGEYNEYLRCYVWAVRTLTITIGLPEPOTLFEIVFOLLNFFSGVFVFSLSLQMRDVTG 446
Db 497 YDGVGNSYRCYVWAVRTLTITIGLPEPOTLFEIVFOLLNFFSGVFVFSLSLQMRDVTG 556
QY 447 AATANQYPRACNDTTIAYNNYSIPKLQKRVRTWYEVTSORSMLDESLLKTLPTTV 506
Db 557 AATAGQTYRSCMDSTVYKYNFYKIPRSVQNRVKTWYEVTSORSMLDESLLKTLPTTV 616
QY 507 QLALAIADVNFISIISKVDLFGKCDTOMIYDMLRLKSVLYLPGDFVCKGGEIGKEMIIKH 566
Db 617 RLDAIDVNYISVSKVALFGCDROMIFDMLKRLSVLYLPGDFVCKGGEIGKEMIIKH 676
QY 567 GEVQVLGGPDGKVLVTLKAGSVFGEISLLAAGGNNRTANVVAHFANLLTLDKKTLOE 626
Db 677 GQVQVLGGPDGKVLVTLKAGSVFGEISLLAAGGNNRTANVVAHFANLLTLDKKTLOE 736
QY 627 ILVHYPDSEIRILMKARVLLKQAKTAEATPPRKDLALLFPKKEETPKLFTLLGGTGKA 686
Db 737 ILVHYPDSEIRILMKARVLLKQAKTAEATPPRKDLALLFPKKEETPKLFTLLGGTGKA 790
QY 687 SL-----ARLLKLRQAAQKENSSEEGEENEDKOKENEDKOKENEDK 734
Db 791 GAKGGRGRLALLRLKELALEAAARQOOL-----LEQAKSDEDAVGE-----G 838
QY 735 KENEDKDKGREPEKPLDRPECTA-----SPIAVEEPPHSVRRTVLP 776
Db 839 SASPEQPPRPEPAPEAPEPTAPEPAPEAPEAPEAPEAPEAPEAPEAPEAPEAPEAPEAPE 898
QY 777 RGTSRQSLIISMAPSAGGGEVLTIEVKEKAKQ 809
Db 899 ---EEHPVRIHVTGLGDPDPSEQILLVEVPEKQEE 928

RESULT 13
CNGB1_BOVIN STANDARD; PRT; 1394 AA.
ID CNGB1_BOVIN STANDARD; PRT; 1394 AA.
AC Q28181; Q28181; Q28082;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 240 kDa protein of rod photoreceptor CNG-channel (Contains: Glutamic
DE acid-rich protein (GARP); Cyclic nucleotide-gated cation channel 4
DE (CNG channel 4) (CNG-4) (Cyclic nucleotide-gated cation channel
```

```

DE modulatory subunit)).
GN Name=CNGB1; Synonyms=CNGC4;
OS Mus taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RX MEDLINE=96009859; PubMed=7546742; DOI=10.1016/0896-6273(95)90151-5;
RA Koerschen H.G., Illing M., Ludwig A., Sautter A., Hofmann F.;
RA Gotzes S., Colville C., Mueller F., Dose A., Godde M., Molday L.,
RA Kaupp U.B., Molday R.S.;
RT "A 240 kDa protein represents the complete beta subunit of the cyclic
RT nucleotide-gated channel from rod photoreceptor.";
RL Neuron 15:627-636(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 454-1394 (ISOFORMS CNG4C; CNG4D AND CNG4E).
RC TISSUE=retina;
RX MEDLINE=96198098; PubMed=8626431; DOI=10.1074/jbc.271.11.6349;
RA Biel M., Zong X., Ludwig A., Sautter A., Hofmann F.;
RT "Molecular cloning and expression of the modulatory subunit of the
RT cyclic nucleotide-gated cation channel.";
RL J. Biol. Chem. 271:6349-6355(1996).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1-590.
RC TISSUE=Retina;
RX PubMed=2014230;
RA Sugimoto Y., Yatsunami K., Tsujimoto M., Khorana H.G., Ichikawa A.;
RT "The amino acid sequence of a glutamic acid-rich protein from bovine
RT retina as deduced from the cDNA sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3116-3119(1991).
CC -!- SUBUNIT: Forms functional heterooligomeric channels with CNG3.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Comment-Isoform CNG4D is the most frequent isoform
CC (CNG4D:CNG4C:CNG4E = 20:2:1) in testis;
CC Name=CNG4C;
CC IsoId=Q28181-1; Sequence=Displayed;
CC Name=CNG4D;
CC IsoId=Q28181-2; Sequence=VSP_001109;
CC Name=CNG4E;
CC IsoId=Q28181-3; Sequence=VSP_001108;
CC -!- TISSUE SPECIFICITY: Retina, testis, kidney, heart and brain.
CC -!- SIMILARITY: Belongs to the cyclic nucleotide-gated cation channel
CC (TC 1.A.1.5) family.
CC -!- SIMILARITY: Contains 1 cyclic nucleotide-binding domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X89626; CAAG1769.1; -; mRNA.
DR EMBL; X94707; CAAG64367.1; -; mRNA.
DR EMBL; M61185; AAA30536.1; -; mRNA.
DR PIR; A40437; A40437.
DR InterPro; IPR000595; CNGP bd.
DR InterPro; IPR005821; Ion_trans.
DR Pfam; PF00027; CNGP_binding.1.
DR Pfam; PF0520; Ion_trans.1.
DR PROSITE; PS00888; CNGP_BINDING_1; 1.
DR PROSITE; PS00889; CNGP_BINDING_2; 1.
DR PROSITE; PS00442; CNGP_BINDING_3; 1.
KW Alternative splicing; CAMP; cAMP-binding; Direct protein sequencing;
KW Glycoprotein; Ion transport; Ionic channel; Multigene family;
KW Nucleotide-binding; Transmembrane; Transport.
FT CHAIN 1 590 Glutamic acid-rich protein.
FT CHAIN 454 1394 Cyclic-nucleotide-gated cation channel 4.
FT TOPO_DOM 1 770 Cytoplasmic (Potential).
FT TRANSMEM 771 789 HI (Potential).

```

---

```

FT TOPO_DOM 790 803 Extracellular (Potential).
FT TRANSMEM 804 822 H2 (Potential).
FT TOPO_DOM 823 847 Cytoplasmic (Potential).
FT TRANSMEM 848 867 H3 (Potential).
FT TOPO_DOM 868 904 Extracellular (Potential).
FT TRANSMEM 905 927 H4 (Potential).
FT TOPO_DOM 928 971 Cytoplasmic (Potential).
FT TRANSMEM 972 991 H5 (Potential).
FT TOPO_DOM 992 1075 Extracellular (Potential).
FT TRANSMEM 1076 1096 H6 (Potential).
FT TOPO_DOM 1097 1222 Cytoplasmic (Potential).
FT NP_BIND 1223 1283 CAMP (By similarity).
FT BINDING 1144 1156 CAMP (Potential).
FT BINDING 1156 1156 CAMP (Potential).
FT CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).
FT VARSPLIC 515 532 Missing (in isoform CNG4E).
FT VARSPLIC 522 530 Missing (in isoform CNG4D).
FT CONFLICT 341 341 /FTID=VSP_001109.
FT CONFLICT 454 465 REEEDDEEED -> MRAGQKGR (in Ref. 2).
FT CONFLICT 482 482 R -> Q (in Ref. 2 and 3).
FT CONFLICT 499 499 A -> T (in Ref. 3).
FT CONFLICT 572 590 VPATEEHEPELOVEDADADS -> GSFQMSPFALORCEALAK
FT CONFLICT 1283 1283 R (in Ref. 3).
FT CONFLICT 1289 1289 S -> A (in Ref. 2).
FT CONFLICT 1336 1336 R -> A (in Ref. 2).
FT CONFLICT 1338 1338 D -> E (in Ref. 2).
FT CONFLICT 1394 AA; 155065 MW; EE6DA59BE3744A7 CRC64;
SQ SEQUENCE 1394 AA; 155065 MW; EE6DA59BE3744A7 CRC64;

Query Match 45.3%; Score 1920; DB 1; Length 1394;
Best Local Similarity 45.9%; Pred. No. 3.2e-103;
Matches 401; Conservative 118; Mismatches 24; Indels 112; Gaps 17.

QY 15 GENNEEQRNRNEEGSHPSNQSOQTAAEQENKGEKSLKTKSTPVTSEEPHTNODK 72
DB 532 GAQAQGEVGAQBQDQGVGA-QDQSTSHQELQ--EEALADSSGVPAETEEHPELOVEDADA 588
QY 73 -----LSKNSGDLTTNPDPQNAAEPTGTVP-----EOKEMDPCK 108
DB 589 DSRPLIAEENPPSPVQLPLSP--AKSDTLAVPGSATGSLRKLRLPSQDDEAEELKMLSPAA 646
QY 109 -----EGPNSPQNKPPA-----APVNEYADAQLHNLVKRMQRORTALYKKKLVE 152
DB 647 SPVVAWSPTSPQGTDDQDRATSTASQNSAIND---RLQELVKLFKERTKVKELID 702
QY 153 GDLS-----SPASPTAKPTAVPPVK--ESDDKPTHEHYRLWFKVKMKPLTEYLKRIKL 206
DB 703 PDVTSDEESPKSPAKKAPAPAPYKPAEQVEEHYCEMLCKCKRPRPKKY----QF 758
QY 207 PMSIDSYTRLYLLWLLAVTAYNNWNCWFILRLVFPYQTADNIHYWLIADIICDIIVLY 266
DB 759 PQSIDPLNLMIYLLWLFVLAWNWNCWLIPIVWAFPYQTPTDNIHLWLLMDYLCCLIVLL 818
QY 267 DMLFTQPRLQVRGGDIIVDSNELRKHRTSTKTFOLDVASIIPEDICYLFFGFNPMFRAN 126
DB 819 DITVFQMLQFVRGGDIITDKKEMRNNVKVSORFKMDMLCLLPDLVLLKFGVNPPLRLP 878
QY 327 RMLKYSFFENHLESIMDKAYIYRVITRTGYLLFILHINACVYVWASNYEGITTRWV 386
DB 879 RCLYMAFFENNRLESILSKAYVYVIRTTAYLLSLHNSCLYVWASAYEGLGSTHWV 938
QY 387 YDGEQNEVLRVYAVRVLTIIGGLPEPQTLFEIVFQLNFFSGVFVFSLLIGQMRDVG 446
DB 939 YDGVGNSVIRCYVAVKTLTIIGGLPDPTLFEIVFQGLNYFTGVFAFSVMIGQMRDVVG 998
QY 447 AATANQNYFRACMDDTIAYMNNYSIPKLQKRVRTWYETWDSQRMLESDDLKTLPTTV 506
DB 999 AATAQTYRSCMDSTVKYMFYKIPRSVQNRVKRTWYETWHSQGMLESSELMVQLPKDM 1058
QY 507 QLALADYNFSLISKVDLFCGCDTQMIYDMLRLKSLVLYLPDGFVCKKGEGEKENYLIKH 566
DB 566

```

Db 1059 RLDAIDVNYISVSKVALFQGCORQIMFDMLKRLRSVYVLPNDYVCKGSEIGREMYIIQA 1118  
 QY 567 GEVQVLGGPDGKVLVTLTKAGSVFGSISLLAAGGNNRRTANVVAHFANLLTLDKKTLOE 626  
 Db 1119 GQVQVLGGPDGKSVLVTLKAGSVFGSISLLAAGGNNRRTANVVAHFANLLTLDKKTLOE 1178  
 QY 627 ILVHPDPSRILMKKARVLLKQKATAEATPRKDLALLFPKKEPPLKFTLLGGTGKA 686  
 Db 1179 ILVHPESQKLLRKARRLNKK-----PKESVLPLPRAGFPKPLFNALAAAGKM 1232  
 QY 687 SL-----ARLLKREQAOKKENSEGEGEKENEDKOKENEDKOKENEDKG 734  
 Db 1233 GAKGGRGRLALLRLKELAAEAARQOOL---LEQAKSDEDAVGE-----G 1280  
 QY 735 KENEDKKGREPEKELDRPECTA-----SPIAVEEPHSVVRTVLP 776  
 Db 1281 SASPEQPRPEPAPEAPAPETAPTEPLAPEAPAPASPPASQERPEGDKDAARP 1340  
 QY 777 RGTSLIISMAPSAGGEEVLTIEVKEKAKQ 809  
 Db 1341 ---EEHPVRIHVTGLGPDSEQILLVPEKQEE 1370

## RESULT 14

Q4SKK9\_TETNG  
 ID Q4SKK9\_TETNG PRELIMINARY; PRT; 573 AA.  
 AC Q4SKK9;  
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)  
 DE Chromosome undetermined SCAF14565, whole genome shotgun sequence.  
 DE (fragment).  
 GN ORFNames=GSTENG00016673001;  
 OS Tetraodon nigroviridis (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 OX NCBI\_TaxID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
 RA Mauceli E., Bounau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
 RA Biemont C., Skalli Z., Cattolico L., Poulin J., De Berardinis V.,  
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,  
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,  
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,  
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.  
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
 the early vertebrate proto-karyotype."  
 RL Nature 431:946-957(2004).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RG Genoscope; Whitehead Institute Centre for Genome Research;  
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; CAAE01014565; CAF98823.1; -; Genomic\_DNA.  
 DR InterPro; IPR005955; cNMP\_bd.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR Pfam; PF00520; cNMP\_binding; 1.  
 DR SMART; SM00100; cNMP; 1.  
 DR PROSITE; PS00888; cNMP\_BINDING\_1; UNKNOWN\_1.  
 DR PROSITE; PS00888; cNMP\_BINDING\_2; UNKNOWN\_1.  
 DR PROSITE; PS00888; cNMP\_BINDING\_3; 1.  
 KW Ion transport; Ionic channel; Transmembrane; Transport.

FT NON\_TER 1 1  
 FT NON\_TER 573 573  
 SQ SEQUENCE 573 AA; 64995 MW; 942B58E39B630BE9 CRC64;  
 Query Match 38.7%; Score 1639.5; DB 2; Length 573;  
 Best Local Similarity 56.4%; Pred. No. 2.3e-87;  
 Matches 326; Conservative 84; Mismatches 149; Indels 19; Gaps 7;  
 QY 118 PPAAPVINEYADAQLHNLVKRQRTALYKKKLVEGDLSLSPASPTAKPTAVPPVKESD 177  
 Db 1 PPPPVINRYSDEQLRTIKRMRLRLQICKVADQYASSPVPVQEOAKERKSVK 60  
 QY 178 DKPETHYRLWFKVKOM--PLTEYLKRIKLPNSIDSYTDRLYLLWLLVTLAYNNCNWF 235  
 Db 61 ARLTEAY----WSAVDALLEPLEKANDSV-VGVTIDPFTDRRYIAWLSLVTLAFNVTWF 115  
 QY 236 IFLRLVFPYQTADNIHYMLIADIICDIYLYDMLFIQRLQFVRGGDIIVDSNELRKHVR 295  
 Db 116 IARLCFPYHSPGAVPLMSLDLADLVLTLSLVFQPRKQFVKAGDIKDRVMSKKNYR 175  
 QY 296 TSTKFO-LDVASIIIPFDICYLFFGNPMFRANRMLKYTSFFFNHLESIMDKAYIRVI 354  
 Db 176 ESERFKVLDWALLPFDLLYLQFGPKSIIPRANRLKADAFEFSDRLESIMAKAYIRVI 235  
 QY 355 RTTGVLILFILHINACVYVWASNYEGITTRWVYDGEENEYLCYYWAVRTLTITIGLPEP 414  
 Db 236 RTIGVLLFMLHNLNACLYVVASDYQIGLTKWYSGDSAYLCYYFAVRSLLINIGGLNEP 295  
 QY 415 QTLFEIVFOLLNFFSGVFFVSSLIIGOMRDVIGAATANQNYFRACMDDTIAYMNNYSI 474  
 Db 296 HTVFEISFQMTNFTGTVFVSSLIIGOMRDVIGAATAGAYFRSNMONTVAYMVTNRIPS 355  
 QY 475 VQKRVRTWYETWDSQRM-----DESLLKLTPTTVQALALADVNFSIISKVDLPK 526  
 Db 356 VQNRVRTWYTWDAQGLQSHAWYQDESELDDKMLVMTAIAVIDINLATPKIDLPK 415  
 QY 527 GCDTQMIYDMLRLKSVLYLPGDFVCKGCEIGKEMVIIKHGEVOVLGGDGTQVLVTLKA 586  
 Db 416 GCDQQLVDMRLRLKSVLYLPGDFVVKYKGDIGKEMVIIKSGAVOVVGGPDNSIIFVTLKA 475  
 QY 587 GSVFGEISLLAAG--CGNRRTANVVAHFANLLTLDKKTLOEILHYVPOSERILMKKARV 644  
 Db 476 GCVFGEISLLQSKDGNGRNTANVKAHFANLVLEKDLFDILVHYPSQVKLARKGRK 535  
 QY 645 LKQKA-KTAAEATPPRKDLALLFPKKEPPLKFTLLG 681  
 Db 536 LMKAGPAAAKVEERKKGLALFGPKPPTPKLLRAFGG 573  
 RESULT 15  
 Q4T4Q5\_TETNG  
 ID Q4T4Q5\_TETNG PRELIMINARY; PRT; 743 AA.  
 AC Q4T4Q5;  
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)  
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)  
 DE Chromosome undetermined SCAF9565, whole genome shotgun sequence.  
 DE (fragment).  
 GN ORFNames=GSTENG00007198001;  
 OS Tetraodon nigroviridis (Green puffer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 OC Tetraodontidae; Tetraodontidae; Tetraodon.  
 OX NCBI\_TaxID=99883;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,  
 RA Mauceli E., Bounau L., Fischer C., Ozouf-Costaz C., Bernot A.,  
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,  
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,  
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,  
 RA Biemont C., Skalli Z., Cattolico L., Poulin J., De Berardinis V.,

RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,  
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,  
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,  
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,  
RA Laudet V., Schachter V., Quetier F., Saurin F., Scarpelli C.,  
RA Winkler P., Lander E.S., Weissenbach J., Roest Crolious H.,  
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals  
the early vertebrate proto-karyotype.",  
RL Nature 431:946-957(2004).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RG Genoscope; Whitehead Institute Centre for Genome Research;  
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; CAAE01009565; CAF92127.1; -; Genomic\_DNA.  
DR InterPro; IPR005821; Ion trans.  
DR Pfam; PF00520; Ion trans; 1  
KW Ion transport; Ionic channel; Transmembrane; Transport.  
FT NON\_TER 1  
FT NON\_TER 743 743  
SQ SEQUENCE 743 AA; 83811 MW; 27350BFD6A8D226D CRC64;  
  
Query Match 31.3%; Score 1327; DB 2; Length 743;  
Best Local Similarity 40.4%; Pred. No. 5.7e-69;  
Matches 277; Conservative 100; Mismatches 186; Indels 122; Gaps 11;  
  
QY 120 AAPVINEYADAOQLHNLVKRMQRTALYKKLVKGLSSPE----- 159  
DB 5 AASVGSVLVQDLRLNQLVSLFKGTERQKERLVDPDESEADPSACMGPDVDSVTRLQRSKV 64  
  
QY 160 -----ASPOTAKPTAVPPVKESDDKPTHE-----YYRLLMFKVK 193  
DB 65 NQLGEALIVGFLISAAPSKAAPPPPPPPPPGDKDEAPAAQAQEEDEPELPFKILGRPVK 124  
  
QY 194 ----KMPLELYLXRI---KLPNSIDSVTRDLYLLWLLVTLAYNWCWFPIPLRVFPYQTA 247  
DB 125 LPRPPLPAWIRLVMEFRFPTSDPYDYYVWLFVVAANWNNVWLIIPVMAFPYQTP 184  
  
QY 248 DNIHWYLIADIIIDIYLDMLFIQPRQFVRGGDIIV----- 285  
DB 185 DNIHLWLLADYCDLIYIADILLFQTRQFVRGGDIVGQTDQLRLAFSSASNRVLSFQ 244  
  
QY 286 -DSNELKHYRTSTKPLQDVASIIIPDICYLFFGNPMFRANMLKYTSFFFNHLES 344  
DB 245 CDRKEMRENYMTDRFKTDLLSLPLELCYIFGVNPLRFPRLKLYQVFFEFNDRMEAV 304  
  
QY 345 MDKAYIYRVIRTTGYLLFTLHINACVYVWASNYEGITTRVYDGEENEYLRCCYYWVRT 404  
DB 305 MKKAYIYRVIRTSYLLSLHINACLFYWGSAYEGLGATKWYDGGKNAYIRCYFVAVKT 364  
  
QY 405 LITIGLPEPQTLFEIVFOLLNFFSGVFVSSLIQMRDVI GAATANQNYFRACMDDTIA 464  
DB 365 LITIGLDPPTTVFELCFQILNYFVGVFAPFSIMIGQMRDVGAANTAGENYRACMDSTVK 424  
  
QY 465 YNNYSIPKLQKRVRTWYETWDSQRMDESLLKTLPTTVQLALADWNFSIISKVDL 524  
DB 425 YNNSYIPIQEVQNRKIKTWYDTWKSQMLDEQLLVQLPTKMHLDIADVNTIVSRVAL 484  
  
QY 525 PKGCDTQMIYDMLLRLKSVLYLPGDFVCKKGKIGKEMWIIKHGEVQVLGGPDGTVLVT 584  
DB 485 FQGCQRQWVDFMLTRLKSVYLLPGDFVCKKVKVTHTHSHSHTLLQ-----RSCNIFCV 539  
  
QY 585 KAGSVFGEISLLAA---GGNRRRTANVAHGFANLLTLDKTLQELVHYHYPDSERILMKK 641  
DB 540 SGRDRQGDVHQAGRGPGGWSRP-----PDLRHQSW 573  
  
QY 642 ARVLLKQKAKTAEATP---PRKDLALLFP--PKEETPKLFTLLGGTGKASLARLLKLKR 696  
DB 574 LVRVGDQLACGRRQPTHQREGSRLRQVHPQGEF-----GGDPGLPRVPEAPP 625  
  
QY 697 EQAAQKENSEGEEGKEDKQK 721

Db 636 QEGQDDADEGQEARTEGRRORDERR 650  
Search completed: December 21, 2005, 19:58:05  
Job time : 238 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2005, 19:46:13 ; Search time 186 Seconds  
(without alignments)  
1911.061 Million cell updates/sec

Title: US-09-855-828-1  
Perfect score: 4234  
Sequence: 1 MFKSTKVNKVPKIGNNEN.....PSAEGGEVLTIEVKEKAKQ 809

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*  
9: Geneseqp2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4234	100.0	809	5	Aae15982 Human cyc
2	4231	99.9	809	5	Aae15983 Human CNG
3	4231	99.9	809	5	Aae15985 Human CNG
4	4229	99.9	809	5	Aae15984 Human CNG
5	4228	99.9	809	5	Aae15986 Human CNG
6	3217	76.0	615	9	Adz88289 Human CNG
7	1952	46.1	1245	9	Adz88287 Human CNG
8	1949	46.0	909	9	Aeal17267 Human cyc
9	1948.5	46.0	1250	8	Abm83979 Human dia
10	1781.5	42.1	652	4	Abg27471 Novel hum
11	1103	26.1	1037	4	Abb66545 Drosophil
12	1088	25.7	644	4	Abg05466 Novel hum
13	1052.5	24.9	821	9	Adz88294 Human CNG
14	955	22.6	747	4	Abg11969 Novel hum
15	906	21.4	237	6	Abg13936 Amino aci
16	902	21.3	237	6	Abp98584 Amino aci
17	850	20.1	694	5	Aae15987 Human CNG
18	850	20.1	694	7	Aae38591 Human CNG
19	850	20.1	694	9	Adz88286 Human CNG
20	831.5	19.6	683	7	Add48638 Rat Prote
21	829	19.6	663	5	Aag79527 Cow HBMYC
22	829	19.6	663	6	Abg74912 Bovine CN
23	829	19.6	663	9	Adz88292 Bull CNG
24	826.5	19.5	698	8	Adq67650 Novel hum

25	825.5	19.5	664	5	AAG79525	Variant H
26	824	19.5	663	6	ABG74914	Abg74914 Bovine CN
27	823	19.4	663	6	ABG74913	Abg74913 Bovine CN
28	821.5	19.4	664	6	ABG72530	Abg72530 Novel hum
29	820.5	19.4	664	5	AAG79524	Aag79524 HBMYCNG.
30	820	19.4	732	5	AAG79526	Aag79526 Rabbit HB
31	820	19.4	732	9	AD288291	Ad288291 Rabbit CN
32	819.5	19.4	664	4	AAE04894	Aae04894 Human tre
33	819.5	19.4	664	5	ABF78066	Abf78066 Amino aci
34	819.5	19.4	664	6	ABG72529	Abg72529 Novel hum
35	819.5	19.4	664	6	ABU12049	Abu12049 Human NOV
36	819.5	19.4	664	6	ABP98475	Abp98475 Amino aci
37	819.5	19.4	664	9	AD288285	Ad288285 Human CNG
38	818.5	19.3	690	7	ADD93244	Add93244 Cyclic nu
39	818.5	19.3	690	8	ADI38341	Adi38341 Human pro
40	818.5	19.3	690	9	AD288284	Ad288284 Human CNG
41	817	19.3	686	7	ADD48640	Add48640 Human pro
42	816.5	19.3	664	8	ADR09903	Adr09903 Human pro
43	815	19.2	690	5	AAE15988	Aae15988 Human CNG
44	805.5	19.0	664	9	AD288293	Ad288293 CNG chann
45	804	19.0	664	5	AAG79529	Aag79529 Rat HBMYC

ALIGNMENTS

RESULT 1	AAE15982					
ID	AAE15982	standard; protein; 809 AA.				
XX	AAE15982;					
AC	AAE15982;					
XX						
DT	26-MAR-2002	(first entry)				
XX						
DE	Human cyclic nucleotide-gated cation channel 3 beta subunit protein.					
XX						
KW	Human; cyclic nucleotide-gated cation channel 3 beta subunit; CNG3B;					
KW	Cyclic nucleotide-gated ion channel; contraceptive; vision disorder;					
KW	male infertility; genetic defect; reporter-ligand interaction; CNG;					
KW	viral infection; cancer.					
XX						
OS	Homo sapiens.					
XX						
FH	Key	Location/Qualifiers				
FT	Region	210..661				
FT		/note= "Conserved region; this region also function as an				
FT		epitope and is referred in claim 1"				
XX						
PN	WO200188090-A2.					
XX						
PD	22-NOV-2001.					
XX						
PF	15-MAY-2001; 2001WO-US015814.					
XX						
PR	15-MAY-2000; 2000US-0204445P.					
PR	14-MAY-2001; 2001US-00859828.					
XX	(ICAG-) ICAGEN INC.					
PA						
XX						
PI	Crech CD, Jegla TJ;					
XX						
DR	WPI; 2002-089847/12.					
XX	N-PSDB; AAD25729.					
DR						
XX						
PT	New polypeptide, useful for screening for modulators of cyclic nucleotide-gated ion channels, comprises the isolated cyclic nucleotide-gated cation channel 3 beta subunit.					
FT						
PS	Claim 12; Fig 4; 83pp; English.					
XX						
CC	The invention relates to human cyclic nucleotide-gated cation channel 3 beta subunit (CNG3B) polypeptides and polynucleotides. CNG3B is a member of cyclic nucleotide-gated cation channel (CNG) family. CNG3B forms					

CC functional heteromultimers with CNGA1 and/or CNGB3, two CNG alpha  
CC subunits that are expressed in retina. CNG3B polypeptides are useful for  
CC screening modulators of CNGs which are useful as contraceptives and for  
CC treating various disorders involving cation channels, e.g. vision  
CC disorders and male infertility. Polynucleotides of the invention are  
CC useful for transfection of cells in vitro and in vivo, to correct  
CC acquired and inherited genetic defects, cancer and viral infections.  
CC Sequences of the invention are useful as reporter molecules in assays and  
CC detection systems, to measure changes in cation concentration, membrane  
CC potential, current flow, ion flux, transcription, signal transduction,  
CC reporter-ligand interactions and second messenger concentrations, in  
CC vitro, in vivo and ex vivo. They are useful to construct models of CNGs  
CC in a computer system and for examining expression and regulation of  
CC cation channels. The present sequence is human CNG3B protein  
XX  
XX Sequence 809 AA;

Query Match 100.0%; Score 4234; DB 5; Length 809;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 809; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFKSLTKVNVKPIGNNENEGSSRRENGSHPSNOSQQTAAQENKGEKSLTKSTPV 60  
Db 1 MFKSLTKVNVKPIGNNENEGSSRRENGSHPSNOSQQTAAQENKGEKSLTKSTPV 60  
QY 61 TSEEPHTNIQDKLSKKNSSGDLTTNPDQNAAEPTGTVPQEMDPGKGNPSQNKPPA 120  
Db 61 TSEEPHTNIQDKLSKKNSSGDLTTNPDQNAAEPTGTVPQEMDPGKGNPSQNKPPA 120  
QY 121 APVINEYADAQLHNLVKNRQRTALYKKLVGDLSSPEASQPTAKPTAVPVKGSDDKP 180  
Db 121 APVINEYADAQLHNLVKNRQRTALYKKLVGDLSSPEASQPTAKPTAVPVKGSDDKP 180  
QY 181 TEHYRLLWFKVKMPLTEYLKRIKLPSNIDSVTDRLYLMLLLVTLAYNMCWFIPRL 240  
Db 181 TEHYRLLWFKVKMPLTEYLKRIKLPSNIDSVTDRLYLMLLLVTLAYNMCWFIPRL 240  
QY 241 VFPYQTADNIHWLIADIICDIYLYDMLFIQPLQFVFGGDIIVDSNELRKHRTSTKF 300  
Db 241 VFPYQTADNIHWLIADIICDIYLYDMLFIQPLQFVFGGDIIVDSNELRKHRTSTKF 300  
QY 301 QLDVASIIFPDCYLPFGNPNFRANMLKYTSFFEFNHLESIMDKAYIYVIRTTGYL 360  
Db 301 QLDVASIIFPDCYLPFGNPNFRANMLKYTSFFEFNHLESIMDKAYIYVIRTTGYL 360  
QY 361 LFILHNACVYVWASNYEGITRWYVYDGEHNEYLRCYVWVRTLITIGLPEPOTLREI 420  
Db 361 LFILHNACVYVWASNYEGITRWYVYDGEHNEYLRCYVWVRTLITIGLPEPOTLREI 420  
QY 421 VFQLNFFSGVVFSSLIQOMRDVIGAAATANQNYFRACMDDTIAYNNYSIPKLQKRV 480  
Db 421 VFQLNFFSGVVFSSLIQOMRDVIGAAATANQNYFRACMDDTIAYNNYSIPKLQKRV 480  
QY 481 TWYETWDSQRMLESDDLKTLPTTVQLALADVNFISIISKVDLFGKCDTQIMYDMLRL 540  
Db 481 TWYETWDSQRMLESDDLKTLPTTVQLALADVNFISIISKVDLFGKCDTQIMYDMLRL 540  
QY 541 KSVLYLPQDFVCKKGIGKEMWIIKHGEVQVVGPDGTVLTKAGSVFGISLLAAG 600  
Db 541 KSVLYLPQDFVCKKGIGKEMWIIKHGEVQVVGPDGTVLTKAGSVFGISLLAAG 600  
QY 601 GNRRTANVVAHFANLLTLDKKTQLQBIYVHPDSEIRLMKKARVLLKQAKTAEATPPRK 660  
Db 601 GNRRTANVVAHFANLLTLDKKTQLQBIYVHPDSEIRLMKKARVLLKQAKTAEATPPRK 660  
QY 661 DLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREQAQKENSEGEGEGKKNEDKQ 720  
Db 661 DLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREQAQKENSEGEGEGKKNEDKQ 720  
QY 721 KENEDQKKNEDKKNEDKDKGREPEEKPLDRPECTASPIAVEEPHVSVRTVLPRGTS 780  
Db 721 KENEDQKKNEDKKNEDKDKGREPEEKPLDRPECTASPIAVEEPHVSVRTVLPRGTS 780

QY 781 RQSLIISMAPSAEGGEEVLTTIEVKEKAKQ 809  
Db 781 RQSLIISMAPSAEGGEEVLTTIEVKEKAKQ 809  
RESULT 2  
AAE15983  
ID AAE15983 standard; protein; 809 AA.  
XX  
AC AAE15983;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Human CNG3B protein variant #1 (R142K).  
XX  
KW Human; cyclic nucleotide-gated cation channel 3 beta subunit; CNG3B;  
KW cyclic nucleotide-gated ion channel; contraceptive; vision disorder;  
KW male infertility; genetic defect; reporter-ligand interaction; CNG;  
KW viral infection; cancer; mutant; mutein; variant.  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 142 /note= "Wild type Arg substituted with Lys"  
XX  
PN WO200188090-A2.  
XX  
PD 22-NOV-2001.  
XX  
PF 15-MAY-2001; 2001WO-US015814.  
XX  
PR 15-MAY-2000; 2000US-0204445P.  
PR 14-MAY-2001; 2001US-00855828.  
XX  
PA (ICAG-) ICAGEN INC.  
PI Creech CD, Jegla TJ;  
XX  
DR WPI; 2002-089847/12.  
XX  
PT New polypeptide, useful for screening for modulators of cyclic nucleotide-  
PT gated ion channels, comprises the isolated cyclic nucleotide gated  
PT cation channel 3 beta subunit.  
XX  
PS Disclosure; Page; 83pp; English.

CC The invention relates to human cyclic nucleotide-gated cation channel 3  
CC beta subunit (CNG3B) polypeptides and polynucleotides. CNG3B is a member  
CC of cyclic nucleotide-gated cation channel (CNG) family. CNG3B forms  
CC functional heteromultimers with CNGA1 and/or CNGB3, two CNG alpha  
CC subunits that are expressed in retina. CNG3B polypeptides are useful for  
CC screening modulators of CNGs which are useful as contraceptives and for  
CC treating various disorders involving cation channels, e.g. vision  
CC disorders and male infertility. Polynucleotides of the invention are  
CC useful for transfection of cells in vitro and in vivo, to correct  
CC acquired and inherited genetic defects, cancer and viral infections.  
CC Sequences of the invention are useful as reporter molecules in assays and  
CC detection systems, to measure changes in cation concentration, membrane  
CC potential, current flow, ion flux, transcription, signal transduction,  
CC reporter-ligand interactions and second messenger concentrations, in  
CC vitro, in vivo and ex vivo. They are useful to construct models of CNGs  
CC in a computer system and for examining expression and regulation of  
CC cation channels. The present sequence is human CNG3B protein variant  
CC (R142K). Note: This sequence is not shown in the specification, but is  
CC derived from the human wild type CNG3B protein (SEQ ID NO:1) shown in Fig  
CC 4 of the specification (AAE15982)

SQ Sequence 809 AA;

Query Match 99.9%; Score 4231; DB 5; Length 809;  
Best Local Similarity 99.9%; Pred. No. 0;

Matches	808;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MPKSLTKVNVKVP	IGENNENEGSSRRNEGSHPSNQSQTTAQEENKGEKSLTKSTPV	60					
Db	1	MPKSLTKVNVKVP	IGENNENEGSSRRNEGSHPSNQSQTTAQEENKGEKSLTKSTPV	60					
Qy	61	TSEEPHTNIQDKLSKNSSGDLTTNPDPPQNAAEPTGTVPQEKMDPGKGNPQNKPPA	120						
Db	61	TSEEPHTNIQDKLSKNSSGDLTTNPDPPQNAAEPTGTVPQEKMDPGKGNPQNKPPA	120						
Qy	121	APVINEYADAQLHNLVKRMQRORTALYKKKLVEGDLSSPEASPTAKPTAVPPVKESDDKP	180						
Db	121	APVINEYADAQLHNLVKRMQRORTALYKKKLVEGDLSSPEASPTAKPTAVPPVKESDDKP	180						
Qy	181	TEHYRLLWPKVKMPLTEYLKRIKLPNSIDSYTDRLLYLLWLLVTLAYNWCWF1PLRL	240						
Db	181	TEHYRLLWPKVKMPLTEYLKRIKLPNSIDSYTDRLLYLLWLLVTLAYNWCWF1PLRL	240						
Qy	241	VFPYQTADNIHYWLIADIICDIIYLYDMLFIQRLQFVRGGDIIVDSNELRKHYRTSTKF	300						
Db	241	VFPYQTADNIHYWLIADIICDIIYLYDMLFIQRLQFVRGGDIIVDSNELRKHYRTSTKF	300						
Qy	301	QLDVASIIIPDICYLFGFENPFRANMLKYTSFFFNHLSIMDKAYIRVIRTTGYL	360						
Db	301	QLDVASIIIPDICYLFGFENPFRANMLKYTSFFFNHLSIMDKAYIRVIRTTGYL	360						
Qy	361	LFILHINACYVYWASNYEGITRWYVYDGEENYLCYYWAVRTLITIGGLPEPQTLFEI	420						
Db	361	LFILHINACYVYWASNYEGITRWYVYDGEENYLCYYWAVRTLITIGGLPEPQTLFEI	420						
Qy	421	VPOLLNPFSGVVFSSLIQMRDVGIAATANQNYFRACMDDTIAYMNNYSIPKLQKVR	480						
Db	421	VPOLLNPFSGVVFSSLIQMRDVGIAATANQNYFRACMDDTIAYMNNYSIPKLQKVR	480						
Qy	481	TWTEYTWDSORMLDESLLKTLPTTQALAIQVNFSSIIISKVDLFGKCDTQMIYDMLLRL	540						
Db	481	TWTEYTWDSORMLDESLLKTLPTTQALAIQVNFSSIIISKVDLFGKCDTQMIYDMLLRL	540						
Qy	541	KSVLYLPGDPVCVKKEIGKEMYIKHGEVQVLGGPDGTVLTKAGSVFGEISLLAAGG	600						
Db	541	KSVLYLPGDPVCVKKEIGKEMYIKHGEVQVLGGPDGTVLTKAGSVFGEISLLAAGG	600						
Qy	601	GNRRTANVAHGAFANLLTLDKTLQELVHYPSERILMKKARVLLKQAKTAEATPPRK	660						
Db	601	GNRRTANVAHGAFANLLTLDKTLQELVHYPSERILMKKARVLLKQAKTAEATPPRK	660						
Qy	661	DLALLPPKEETPKLFTLLGGTGKASLARLLKREQAQKKNSESGGEEGKENEKQ	720						
Db	661	DLALLPPKEETPKLFTLLGGTGKASLARLLKREQAQKKNSESGGEEGKENEKQ	720						
Qy	721	KENEDKQENEDKDKGREPEEKPLDRPECTASPIAVEEPPHSVRRTVLPRTGS	780						
Db	721	KENEDKQENEDKDKGREPEEKPLDRPECTASPIAVEEPPHSVRRTVLPRTGS	780						
Qy	781	RQSLIISMAPSAEGGEEVLTVIEVKEKAKQ 809							
Db	781	RQSLIISMAPSAEGGEEVLTVIEVKEKAKQ 809							
RESULT 3									
ID	AAE15985	standard; protein; 809 AA.							
XX	XX								
AC	AC	AAE15985;							
DT	DT	26-MAR-2002 (first entry)							
XX	XX	Human CNG3B protein variant #3 (L675V).							
XX	XX	Human; cyclic nucleotide-gated cation channel 3 beta subunit; CNG3B;							
KW	KW	cyclic nucleotide-gated ion channel; contrareptive; vision disorder;							
KW	KW	male infertility; genetic defect; reporter-ligand interaction; CNG;							
KW	KW	viral infection; cancer; mutant; mutein; variant.							

Db 241 VFPYQTADNIHWLIADIICDIYLYDMLFIQPRQLQFVRGGDIIVDSNELRKHVRTSTKF 300  
QY 301 QLDVASIIPFDICYLFFGFGNPFNRANMLKYTSFFEFNHLESIMDKAYIYRVIRTTGYL 360  
Db 301 QLDVASIIPFDICYLFFGFGNPFNRANMLKYTSFFEFNHLESIMDKAYIYRVIRTTGYL 360  
QY 361 LFIHINACVYVWASNYEGIGTRWYVYDGEYLRCCYYWAVRTLITIGLPEPOTLFEI 420  
Db 361 LFIHINACVYVWASNYEGIGTRWYVYDGEYLRCCYYWAVRTLITIGLPEPOTLFEI 420  
QY 421 VFQLNFFSGVVFSSLIQOMRDVIGAAATANQNYFRACDDTIAYMNNYSIPKLQKRV 480  
Db 421 VFQLNFFSGVVFSSLIQOMRDVIGAAATANQNYFRACDDTIAYMNNYSIPKLQKRV 480  
QY 481 TWYEYTWDSQRMDESLLKTLPTTVQLALADVNFSIISKVDLPKGCOTQIMYDMLLRL 540  
Db 481 TWYEYTWDSQRMDESLLKTLPTTVQLALADVNFSIISKVDLPKGCOTQIMYDMLLRL 540  
QY 541 KSVLYLPGDFVCKGEGIKEMVIIKHGEVQVLGGPDGTVLTLKAGSVFGEISLLAAG 600  
Db 541 KSVLYLPGDFVCKGEGIKEMVIIKHGEVQVLGGPDGTVLTLKAGSVFGEISLLAAG 600  
QY 601 GNRRTANVAHGFANLLTLDKTQLQBIYVHPDSEIRILMKARVLLKQAKTAETPPRK 660  
Db 601 GNRRTANVAHGFANLLTLDKTQLQBIYVHPDSEIRILMKARVLLKQAKTAETPPRK 660  
QY 661 DLALLFPKKEETPKFTLLGGTGKASLARLLKREQAOKKENSEGEGEKENEDKQ 720  
Db 661 DLALLFPKKEETPKFTLLGGTGKASLARLLKREQAOKKENSEGEGEKENEDKQ 720  
QY 721 KENEDKQENEDKGENEDKGRPEEKPLDRPECTASPIAVEBEPHSVRVTLPRGTS 780  
Db 721 KENEDKQENEDKGENEDKGRPEEKPLDRPECTASPIAVEBEPHSVRVTLPRGTS 780  
QY 781 ROSLIISMAPSAGGEEVLTIIEVKEKAKQ 809  
Db 781 ROSLIISMAPSAGGEEVLTIIEVKEKAKQ 809

RESULT 4  
AAE15984  
ID AAE15984 standard; protein; 809 AA.  
XX AC AAE15984;  
DT 26-MAR-2002 (first entry)  
XX DE Human CNG3B protein variant #2 (D154N).  
XX KW Human; cyclic nucleotide-gated cation channel 3 beta subunit; CNG3B;  
KW cyclic nucleotide-gated ion channel; contraceptive; vision disorder;  
KW male infertility; genetic defect; reporter-ligand interaction; CNG;  
KW viral infection; cancer; mutant; mutein; variant.  
XX OS Homo sapiens.  
OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Misc-difference 154 /note= "wild type Asp substituted with Asn"  
XX WO200188090-A2.  
PN 22-NOV-2001.  
PD 15-MAY-2001; 2001WO-US015814.  
XX PF 15-MAY-2000; 2000US-0204445P.  
PR 14-MAY-2001; 2001US-0085828.  
XX PA (ICAG-) ICAGEN INC.

PI Crech CD, Jegla TJ;  
XX WPI; 2002-089847/12.  
XX New polypeptide, useful for screening for modulators of cyclic nucleotide-  
PT gated ion channels, comprises the isolated cyclic nucleotide-gated  
PT cation channel 3 beta subunit.  
XX Disclosure; Page; 83pp; English.  
XX The invention relates to human cyclic nucleotide-gated cation channel 3  
CC beta subunit (CNG3B) polypeptides and polynucleotides. CNG3B is a member  
CC of cyclic nucleotide-gated cation channel (CNG) family. CNG3B forms  
CC functional heteromultimers with CNGA1 and/or CNGB3, two CNG alpha  
CC subunits that are expressed in retina. CNG3B polypeptides are useful for  
CC screening modulators of CNGs which are useful as contraceptives and for  
CC treating various disorders involving cation channels, e.g. vision  
CC disorders and male infertility. Polynucleotides of the invention are  
CC useful for transfection of cells in vitro and in vivo, to correct  
CC acquired and inherited genetic defects, cancer and viral infections.  
CC Sequences of the invention are useful as reporter molecules in assays and  
CC detection systems, to measure changes in cation concentration, membrane  
CC potential, current flow, ion flux, transcription, signal transduction,  
CC reporter-ligand interactions and second messenger concentrations, in  
CC vitro, in vivo and ex vivo. They are useful to construct models of CNGs  
CC in a computer system and for examining expression and regulation of  
CC cation channels. The present sequence is human CNG3B protein variant  
CC (D154N). Note: This sequence is not shown in the specification, but is  
CC derived from the human wild type CNG3B protein (SEQ ID NO:1) shown in Fig  
CC 4 of the specification (AAE15982)  
XX SQ Sequence 809 AA;  
Query Match 99.9%; Score 4229; DB 5; Length 809;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 808; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFKSLTKVNVKPIGNNENEGSSRRNEEGSHPSNQSOOTTAQENKGEESLTKSTPV 60  
Db 1 MFKSLTKVNVKPIGNNENEGSSRRNEEGSHPSNQSOOTTAQENKGEESLTKSTPV 60  
QY 61 TSEEPHTNIQDKLSKKNSSGDLTTNPDQNAAEPTGTVPQEKMDPGKKGPNISQNKTPA 174  
Db 61 TSEEPHTNIQDKLSKKNSSGDLTTNPDQNAAEPTGTVPQEKMDPGKKGPNISQNKTPA 174  
QY 121 APVINEYADAQHNLVKRMORTALYKKKLVEGDLSSPEASPTAKPTAVPVKESDDKP 180  
Db 121 APVINEYADAQHNLVKRMORTALYKKKLVEGDLSSPEASPTAKPTAVPVKESDDKP 180  
QY 181 TEHYRLLWFKVKMPLTEYLKRIKLPSIDSYTDRLYLLMLLVTLAYNNCWFIPURL 240  
Db 181 TEHYRLLWFKVKMPLTEYLKRIKLPSIDSYTDRLYLLMLLVTLAYNNCWFIPURL 240  
QY 241 VFPYQTADNIHWLIADIICDIYLYDMLFIQPRQLQFVRGGDIIVDSNELRKHVRTSTKF 300  
Db 241 VFPYQTADNIHWLIADIICDIYLYDMLFIQPRQLQFVRGGDIIVDSNELRKHVRTSTKF 300  
QY 301 QLDVASIIPFDICYLFFGFGNPFNRANMLKYTSFFEFNHLESIMDKAYIYRVIRTTGYL 360  
Db 301 QLDVASIIPFDICYLFFGFGNPFNRANMLKYTSFFEFNHLESIMDKAYIYRVIRTTGYL 360  
QY 361 LFIHINACVYVWASNYEGIGTRWYVYDGEYLRCCYYWAVRTLITIGLPEPOTLFEI 420  
Db 361 LFIHINACVYVWASNYEGIGTRWYVYDGEYLRCCYYWAVRTLITIGLPEPOTLFEI 420  
QY 421 VFQLNFFSGVVFSSLIQOMRDVIGAAATANQNYFRACDDTIAYMNNYSIPKLQKRV 480  
Db 421 VFQLNFFSGVVFSSLIQOMRDVIGAAATANQNYFRACDDTIAYMNNYSIPKLQKRV 480  
QY 481 TWYEYTWDSQRMDESLLKTLPTTVQLALADVNFSIISKVDLPKGCOTQIMYDMLLRL 540  
Db 481 TWYEYTWDSQRMDESLLKTLPTTVQLALADVNFSIISKVDLPKGCOTQIMYDMLLRL 540

QY 541 KSVLYLPDGFVCKKGIGKEMYYIKHGEVQVVGPDGTVLTKAGSVFGEISLLAAGG 600  
 DB 541 KSVLYLPDGFVCKKGIGKEMYYIKHGEVQVVGPDGTVLTKAGSVFGEISLLAAGG 600  
 QY 601 GNRRTANVAHGFANLLTLDKKTQILVHYPDSEIRILMKKARVLLKQAKTAETPPRK 660  
 DB 601 GNRRTANVAHGFANLLTLDKKTQILVHYPDSEIRILMKKARVLLKQAKTAETPPRK 660  
 QY 661 DLALLPPKBEETPKLFTKLLGGTGKASLARLLKREQAQKKEGEGEKENEDKQ 720  
 DB 661 DLALLPPKBEETPKLFTKLLGGTGKASLARLLKREQAQKKEGEGEKENEDKQ 720  
 QY 721 KENEDKQKENEDKQKREPEEKPLDRPCTASPIAVEEPPHSVRRTVLPGRGTS 780  
 DB 721 KENEDKQKENEDKQKREPEEKPLDRPCTASPIAVEEPPHSVRRTVLPGRGTS 780  
 QY 781 RQSLIISMAPSAGGEBEVLTIIEVKEKAKQ 809  
 DB 781 RQSLIISMAPSAGGEBEVLTIIEVKEKAKQ 809

RESULT 5  
 AAEL15986  
 ID AAEL15986 standard; protein; 809 AA.  
 XX AAEL15986;  
 DT 26-MAR-2002 (first entry)  
 DE Human CNG3B protein variant #4 (G682S).  
 XX Human; cyclic nucleotide-gated cation channel 3 beta subunit; CNG3B;  
 KW cyclic nucleotide-gated ion channel; contraceptive; vision disorder;  
 KW male infertility; Genetic defect; reporter-ligand interaction; CNG;  
 KW viral infection; cancer; mutant; mutein; variant.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX Key Location/Qualifiers  
 FH Misc-difference 682 /note= "Wild type Gly substituted with Ser"  
 FT WO200188090-A2.  
 XX 22-NOV-2001.  
 XX 15-MAY-2001; 2001WO-US015814.  
 XX 15-MAY-2000; 2000US-0204445P.  
 PR 14-MAY-2001; 2001US-00855828.  
 XX (ICAGEN-) ICAGEN INC.  
 PA  
 XX Crech CD, Jegla TJ;  
 PI WPI; 2002-089847/12.  
 DR  
 XX New polypeptide, useful for screening for modulators of cyclic nucleotide  
 PT -gated ion channels, comprises the isolated cyclic nucleotide-gated  
 PT cation channel 3 beta subunit.  
 XX  
 PS Disclosure; Page; 83pp; English.  
 XX  
 CC The invention relates to human cyclic nucleotide-gated cation channel 3  
 CC beta subunit (CNG3B) polypeptides and polynucleotides. CNG3B is a member  
 CC of cyclic nucleotide-gated cation channel (CNG) family. CNG3B forms  
 CC functional heteromultimers with CNGA1 and/or CNGB3, two CNG alpha  
 CC subunits that are expressed in retina. CNG3B polypeptides are useful for  
 CC screening modulators of CNGs which are useful as contraceptives and for  
 CC treating various disorders involving cation channels, e.g. vision  
 CC disorders and male infertility. Polynucleotides of the invention are  
 CC useful for transfection of cells in vitro and in vivo, to correct

CC acquired and inherited genetic defects, cancer and viral infections.  
 CC Sequences of the invention are useful as reporter molecules in assays and  
 CC detection systems, to measure changes in cation concentration, membrane  
 CC potential, current flow, ion flux, transcription, signal transduction,  
 CC reporter-ligand interactions and second messenger concentrations, in  
 CC vitro, in vivo and ex vivo. They are useful to construct models of CNGs  
 CC in a computer system and for examining expression and regulation of  
 CC cation channels. The present sequence is human CNG3B protein variant  
 CC (G682S). Note: This sequence is not shown in the specification, but is  
 CC derived from the human wild type CNG3B protein (SEQ ID NO:1) shown in Fig  
 CC 4 of the specification (AAEL15982)  
 XX  
 SQ Sequence 809 AA;  
 Query Match 99.9%; Score 4228; DB 5; Length 809;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 808; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MFKSLTKVNVKVPKIGENNENESSRRENEEGSHPSNOSQOTTAQEEKNKGEEKSLTKSTPV 60  
 DB 1 MFKSLTKVNVKVPKIGENNENESSRRENEEGSHPSNOSQOTTAQEEKNKGEEKSLTKSTPV 60  
 QY 61 TSEEPHTNIQDKLSKKNSSGDLTTNPDPQNAAEPTCTVPEQKEMDPGKGPNSPKNPPA 120  
 DB 61 TSEEPHTNIQDKLSKKNSSGDLTTNPDPQNAAEPTCTVPEQKEMDPGKGPNSPKNPPA 120  
 QY 121 APVINEYADAQLHNLVKRMQRORTALYKKLVKVEGDLSSPEASPOKAPTAVPPVKESDDKP 180  
 DB 121 APVINEYADAQLHNLVKRMQRORTALYKKLVKVEGDLSSPEASPOKAPTAVPPVKESDDKP 180  
 QY 181 TEHYELLAFKVKKMPLETKLIPNSIDSTYDRLLYLLWLLVTLAYNWNCFPLRL 240  
 DB 181 TEHYELLAFKVKKMPLETKLIPNSIDSTYDRLLYLLWLLVTLAYNWNCFPLRL 240  
 QY 241 VFPYQADNHIHYWLIADIIICDIIYLYDMLFIQPRLQFVRGGDIIVDSNELKHYRTSTKF 300  
 DB 241 VFPYQADNHIHYWLIADIIICDIIYLYDMLFIQPRLQFVRGGDIIVDSNELKHYRTSTKF 300  
 QY 301 QLDVASIIPFDICYLFFGPNMFRANRMLKYTSFFEFNHLESIMDKAYIYRVIRTTGYL 360  
 DB 301 QLDVASIIPFDICYLFFGPNMFRANRMLKYTSFFEFNHLESIMDKAYIYRVIRTTGYL 360  
 QY 361 LFTLHINACVYWASNYEGITTRWYDGEENYLCYYWAVRTLITIGGLPBPOTLFEI 420  
 DB 361 LFTLHINACVYWASNYEGITTRWYDGEENYLCYYWAVRTLITIGGLPBPOTLFEI 420  
 QY 421 VFOLLNFFSGVVFSSLIQOMRDVIGAATANQNYFRACMDDTIAYMNNYSIPKLVKRVR 480  
 DB 421 VFOLLNFFSGVVFSSLIQOMRDVIGAATANQNYFRACMDDTIAYMNNYSIPKLVKRVR 480  
 QY 481 TWYEYTWDSQRMLESDDLKTLPTTVQALAIQVNFSSIIISKVDLFKGCDDTOMIYDMLLRL 540  
 DB 481 TWYEYTWDSQRMLESDDLKTLPTTVQALAIQVNFSSIIISKVDLFKGCDDTOMIYDMLLRL 540  
 QY 541 KSVLYLPDGFVCKKGIGKEMYYIKHGEVQVVGPDGTVLTKAGSVFGEISLLAAGG 600  
 DB 541 KSVLYLPDGFVCKKGIGKEMYYIKHGEVQVVGPDGTVLTKAGSVFGEISLLAAGG 600  
 QY 601 GNRRTANVAHGFANLLTLDKKTQILVHYPDSEIRILMKKARVLLKQAKTAETPPRK 660  
 DB 601 GNRRTANVAHGFANLLTLDKKTQILVHYPDSEIRILMKKARVLLKQAKTAETPPRK 660  
 QY 661 DLALLPPKBEETPKLFTKLLGGTGKASLARLLKREQAQKKEGEGEKENEDKQ 720  
 DB 661 DLALLPPKBEETPKLFTKLLGGTGKASLARLLKREQAQKKEGEGEKENEDKQ 720  
 QY 721 KENEDKQKENEDKQKREPEEKPLDRPCTASPIAVEEPPHSVRRTVLPGRGTS 780  
 DB 721 KENEDKQKENEDKQKREPEEKPLDRPCTASPIAVEEPPHSVRRTVLPGRGTS 780  
 QY 781 RQSLIISMAPSAGGEBEVLTIIEVKEKAKQ 809  
 DB 781 RQSLIISMAPSAGGEBEVLTIIEVKEKAKQ 809

RESULT 6  
ADZ88289  
ID ADZ88289 standard; protein; 615 AA.  
XX  
AC ADZ88289;  
XX  
DT 14-JUL-2005 (first entry)  
XX  
DE Human CNG channel CNGB3.  
XX  
KW antiasthmatic; Antiallergic; Hypotensive; Gastrointestinal-Gen.;  
KW Antiparkinsonian; Antidepressant; Neuroleptic; Cardiant; GPCR-Antagonist;  
KW Unspecified GPCR; cyclic nucleotide-gated channel A2; CNCA2; CNG;  
KW mutation; divalent cation-mediated blockage; cation flux;  
KW G-protein coupled receptor; asthma; hypertension; allergic reaction;  
KW Parkinsons disease; depression; schizophrenia; heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO2005041758-A2.  
XX  
PD 12-MAY-2005.  
XX  
PF 01-NOV-2004; 2004WO-US036022.  
XX  
PR 30-OCT-2003; 2003US-0515442P.  
XX  
PR 20-JUL-2004; 2004US-0589012P.  
XX  
PA (ATTO-) ATTO BIOSCIENCE INC.  
XX  
PI Yao Y, Cao L, Lu J, Lorente I;  
XX  
XX WPI; 2005-366496/37.  
XX  
XX New isolated polynucleotide encoding a mutant cyclic nucleotide-gated  
PT (CNG) channel having a mutation that decreases divalent cation-mediated  
PT blockage of cation flux, and makes the channel more sensitive to cAMP.  
XX  
XX Disclosure; SEQ ID NO 14; 156pp; English.  
XX  
XX This sequence represents a wild type human cyclic nucleotide-gated (CNG)  
CC channel. The rat CNG channel sequence, CNCA2, may be mutated to form the  
CC isolated polynucleotide of the invention which comprising a sequence  
CC encoding a mutant CNG channel comprising at least one mutation that  
CC decreases divalent cation-mediated blockage of cation flux, and at least  
CC one mutation that makes the channel more sensitive to cAMP than a channel  
CC that does not comprise the mutation. The invention includes methods for:  
CC (a) detecting changes in intracellular levels of cAMP, comprising  
CC expressing in a cell the polynucleotide cited above, and measuring  
CC activity of the channel, where activity of the channel is indicative of  
CC changes in intracellular cAMP; (b) detecting activity of a GPCR,  
CC comprising expressing the CNG channel and the G-protein coupled receptor  
CC (GPCR) in host cell transformed with the nucleotide of the invention, and  
CC measuring activity of the channel, where activity of the channel  
CC indicates activity of the GPCR; (c) identifying a ligand for a receptor,  
CC comprising contacting a cell with a compound wherein the cell expresses  
CC the receptor and the mutant CNG channel cited above, and measuring  
CC activation of the CNG channel, where activation of the CNG channel  
CC indicates that the compound is a ligand for the receptor; (d) identifying  
CC an agent that modulates an activity mediated by a GPCR, comprising  
CC contacting the transformed host cell with the agent and a ligand for the  
CC receptor, and measuring activation of the CNG channel. The methods and  
CC compositions of the present invention are useful in cellular physiology,  
CC in particular for cell-based assays that employ cation and voltage-based  
CC dyes and other indicators, including novel CNG mutants for cyclic  
CC nucleotide assays, and for measuring the activity of GPCR in disorders  
CC such as asthma, hypertension, allergic reactions, gastrointestinal  
CC disorders, Parkinson's disease, depression, schizophrenia and heart  
CC disease. This sequence was included in the specification for comparison  
CC purposes.

SQ Sequence 615 AA;

Query Match 76.0%; Score 3217; DH 9; Length 615,  
Best Local Similarity 100.0%; Pred. No. 2.4e-265;  
Matches 615; Conservative 0; Mismatches 0; Indels 0; Gaps 0.  
QY 195 MPLTEYLKRIKLPNSIDSYTDRLYLMLLLVLTAYNNWCWFIPRLVFPYQTADNIHWL 254  
DB 1 MPLTEYLKRIKLPNSIDSYTDRLYLMLLLVLTAYNNWCWFIPRLVFPYQTADNIHWL 60  
QY 255 IADIICDIIYLYDMLFIQPRLOFVRGGDIIIVDSNELRKHYRTSTKFOLDVASIIFPDICY 314  
DB 61 IADIICDIIYLYDMLFIQPRLOFVRGGDIIIVDSNELRKHYRTSTKFOLDVASIIFPDICY 120  
QY 315 LFFGPNPFRANRMLKYTSFFEFNHHLESIMDKAYIYRVIRTTGYLLFILHINACVYWA 174  
DB 121 LFFGPNPFRANRMLKYTSFFEFNHHLESIMDKAYIYRVIRTTGYLLFILHINACVYWA 180  
QY 375 SNYEGIGTRWYDGEENEYLRCCYWAVRTLITIGLPEPOTLFEIVFOLLNFFSGVFVF 434  
DB 181 SNYEGIGTRWYDGEENEYLRCCYWAVRTLITIGLPEPOTLFEIVFOLLNFFSGVFVF 240  
QY 435 SSLIGQMRDVI GAATANQNYFRACMDDTIAYNNYSIPKLVQKRVRTWYETWDSQRMLD 494  
DB 241 SSLIGQMRDVI GAATANQNYFRACMDDTIAYNNYSIPKLVQKRVRTWYETWDSQRMLD 300  
QY 495 ESDLLKTLPTTVQLALADIVNFSIISKVDLPFGCDTQMIYDMLLRKSVLYLPGDFVCKK 554  
DB 301 ESDLLKTLPTTVQLALADIVNFSIISKVDLPFGCDTQMIYDMLLRKSVLYLPGDFVCKK 360  
QY 555 GEIGKEMYIIKHGEVQVLGGPDGTVLTKAGSVFGEISLLAAGGNNRRNTANVVAHGFA 614  
DB 361 GEIGKEMYIIKHGEVQVLGGPDGTVLTKAGSVFGEISLLAAGGNNRRNTANVVAHGFA 420  
QY 615 NLLTLDKKTLOEILVHYPDSEIRILMKKARVLLKQAKTAEATPPRKDLALLFPPEETPK 674  
DB 421 NLLTLDKKTLOEILVHYPDSEIRILMKKARVLLKQAKTAEATPPRKDLALLFPPEETPK 480  
QY 675 LFKTLTGGTGKASLARLLKREQAOKKENSEGEEGKENEKOKENEKOKENEKOKENEK 734  
DB 481 LFKTLTGGTGKASLARLLKREQAOKKENSEGEEGKENEKOKENEKOKENEKOKENEK 740  
QY 735 KENEDKGRPEEPKPLDRPECTASPIAVEEPHSVRRTVLPRTGSROSLIISMAPSARG 794  
DB 541 KENEDKGRPEEPKPLDRPECTASPIAVEEPHSVRRTVLPRTGSROSLIISMAPSARG 600  
QY 795 GEEVLTIEVKAKQ 809  
DB 601 GEEVLTIEVKAKQ 615  
RESULT 7  
ADZ88287  
ID ADZ88287 standard; protein; 1245 AA.  
XX  
AC ADZ88287;  
XX  
DT 14-JUL-2005 (first entry)  
XX  
DE Human CNG channel CNGB1.  
XX  
KW antiasthmatic; Antiallergic; Hypotensive; Gastrointestinal-Gen.;  
KW Antiparkinsonian; Antidepressant; Neuroleptic; Cardiant; GPCR-Antagonist;  
KW Unspecified GPCR; cyclic nucleotide-gated channel A2; CNCA2; CNG;  
KW mutation; divalent cation-mediated blockage; cation flux;  
KW G-protein coupled receptor; asthma; hypertension; allergic reaction;  
KW Parkinsons disease; depression; schizophrenia; heart disease.  
XX  
OS Homo sapiens.  
XX  
PN WO2005041758-A2.  
XX  
XX 12-MAY-2005.  
PD

```

XX PF 01-NOV-2004; 2004WO-US036022.
XX PR 30-OCT-2003; 2003US-0515442P.
XX PR 20-JUL-2004; 2004US-0589012P.
XX PA (ATTO-) ATTO BIOSCIENCE INC.
XX PI Yao Y, Cao L, Lu J, Llorente I;
XX WPI; 2005-366456/37.
XX DR New isolated polynucleotide encoding a mutant cyclic nucleotide-gated
XX PT (CNG) channel having a mutation that decreases divalent cation-mediated
XX PT blockage of cation flux, and makes the channel more sensitive to cAMP.
XX PS Disclosure; SEQ ID NO 12; 156pp; English.
XX CC This sequence represents a wild type human cyclic nucleotide-gated (CNG)
XX CC channel. The rat CNG channel sequence, CNGA2, may be mutated to form the
XX CC isolated polynucleotide of the invention which comprising a sequence
XX CC encoding a mutant CNG channel comprising at least one mutation that
XX CC decreases divalent cation-mediated blockage of cation flux, and at least
XX CC one mutation that makes the channel more sensitive to cAMP than a channel
XX CC that does not comprise the mutation. The invention includes methods for:
XX CC (a) detecting changes in intracellular levels of cAMP, comprising
XX CC expressing in a cell the polynucleotide cited above, and measuring
XX CC activity of the channel, where activity of the channel is indicative of
XX CC changes in intracellular cAMP; (b) detecting activity of a GPCR,
XX CC comprising expressing the CNG channel and the G-protein coupled receptor
XX CC (GPCR) in host cell transformed with the nucleotide of the invention, and
XX CC measuring activity of the channel, where activity of the channel
XX CC indicates activity of the GPCR; (c) identifying a ligand for a receptor,
XX CC comprising contacting a cell with a compound wherein the cell expresses
XX CC the receptor and the mutant CNG channel cited above, and measuring
XX CC activation of the CNG channel, where activation of the CNG channel
XX CC indicates that the compound is a ligand for the receptor; (d) identifying
XX CC an agent that modulates an activity mediated by a GPCR, comprising
XX CC contacting the transformed host cell with the agent and a ligand for the
XX CC receptor, and measuring activation of the CNG channel. The methods and
XX CC compositions of the present invention are useful in cellular physiology,
XX CC in particular for cell-based assays that employ calcium and voltage-based
XX CC dyes and other indicators, including novel CNG mutants for cyclic
XX CC nucleotide assays, and for measuring the activity of GPCR in disorders
XX CC such as asthma, hypertension, allergic reactions, gastrointestinal
XX CC disorders, Parkinson's disease, depression, schizophrenia and heart
XX CC disease. This sequence was included in the specification for comparison
XX CC purposes.
XX SQ Sequence 1245 AA;

Query Match 46.1%; Score 1952; DB 9; Length 1245;
Best Local Similarity 46.4%; Pred. No. 8.2e-157;
Matches 405; Conservative 121; Mismatches 237; Indels 110; Gaps 18;

Qy 14 IGENENEQSRNEEGSHFSNOSQQTAAQ--ENKGE---EKSLTKTPTVTSEEPHTN 68
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 403 VGEAKKEAEKAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAE 462
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 69 IQDK-----LSKNSGGLTTNPPDNQNAEPTGTVPQEKEMDPGKEGN----- 112
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 463 VEDTDADSCPLMAENPPS--TVLPPEPSPAKSDTLIVPSSASGTHRXKLPSEDDAEELK 520
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 113 --SPQNKP-----PAAP-----VINEYADAQLNLVKRMORTALY 146
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 521 ALSPAESPVAVMSDPTTKDTGQDRAASTASTNSAIIND-----RLQELVKLFKERTKV 576
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 147 KKKLVEGDLG----SPASQPTAKPTAVPPVKESDDKPT--EHYRLLWFKVKOMPLETY 200
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 577 KEKLIDPDVTSDESPKSPAKPAPEAPDTPKPAEAEPEVEEHYCDMLCKCKFKHPWKY 636
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 201 LKRIKLNSIDSTDRILYLLMLLVLTAYNNWCFIPLRLVFPQTADNIHYWLIADIC 260
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```



CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorders, developmental disorders, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp protein of the  
CC invention. Note: The sequence data for this patent is not represented in  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [www.wipo.int/pct/en/sequences/listing.htm](http://www.wipo.int/pct/en/sequences/listing.htm)  
XX  
SQ Sequence 1250 AA;

Query Match 46.0%; Score 1948.5; DB 8; Length 1250;  
Best Local Similarity 46.2%; Pred. No. 1.6e-156;  
Matches 403; Conservative 125; Mismatches 235; Indels 109; Gaps 18;

QY 14 IGENNEQSRNRNEGSHFSNQSQTQAQ--ENKGE---EKSLTKTKSTPVTSEBPTHN 68  
DB 409 VGEAKKEAEKAE 468  
QY 69 IQDK-----LSKSNSSGDLTTNPDPQNAEPTGTP-----EQKE 103  
DB 469 VEDTDADSCPLMAEENPPS--TVLPSPSPAKSDTLIVPSSASGTHKKLPSEDEAEELKA 526  
QY 104 MDPGK-----EGNSPNQ-----KPPAPVINEYADAOHLNLRMRQRTALYK 147  
DB 527 LSPAESPVAWSDDTPKDTGQDRAASTASTNSAIIND---RLQELVKLFKERTKVK 582  
QY 148 KKLVEGDLG---SPEASPTAKTAVTPPVKESDDKPT--EHYRLLWFKVKMPLTEYL 201  
DB 583 EKLIDPVTSESPKSPAKAPADTPKPAEAEPEVEBEHYCDMLCKCFKHPMKKY- 641  
QY 202 KRIKPLNSIDYDRLLVLLMLLTAYNNWCNFIPLRLVFPYOTADNIHWLIADILCD 261  
DB 642 ---QFQSIDPLTNTMLVLMFFVMAWNNWCNFIPLRVWAFPYQTPDNIIHHLMDLYLD 698  
QY 262 IIVLYDMLFTQPRLOFVRGGDIIIVDSNELRKHRTSTKFDQDVASIIIPDICYLFFGFNP 321  
DB 699 LIYFLDITVQTRLOFVRGGDIIITDKMRNNYLKSRFFKMDLSLPLDPLFLYLKGVNP 758  
QY 322 MFRANRMLKTSFPFNNHLESIMDKAYIVRVIRTTGYLLFILHNACVYVWASNYEGIG 381  
DB 759 LLRLPRCLKYMAFPFNSRLSILSKAYVYRVIRTTAYLLYSLHNSCLLYWASAYQGLG 818  
QY 382 TTRWYDGEYELRYCYWAVRTLTITGGLPEPOTLFEIVFOLLNFPFGVVFSSLLGQM 441  
DB 819 STHWYDGVGNSYIRCYFAVKTLITITGGLPDKTLFEIVFOLLNFTGVFAFVSMIGQM 878  
QY 442 RDVIGATANQNYFRACMDDTIAYNNYSIPKLVQKVRVYETWDSQRMDESLLKT 501  
DB 879 RDVVGATAGTYRSCWDSIVKYNFKIPKSVQNVKVTWYETWHSQGLMDESLLMVQ 938  
QY 502 LPTTVQLALADVNFISIISKVDLPKGDQDTOMIYDMLLRLKSVLPLPGDFVCKGGEIGEM 561  
DB 939 LPDKMLDLADVNVNIVSKVALFGQDQRMIFDMLKRLRSVWVLPNDYVCKGGEIGEM 998  
QY 562 YIIKHGEVQVGGDGTIVLTKAGSVFGISILLAGGNGRRTANVVAHGFANLLTLDK 621  
DB 999 YIIQAGVQVGGDGTIVLTKAGSVFGISILLAGGNGRRTANVVAHGFANLLTLDK 1058  
QY 622 KTLQELVHYDPSRILMKKARVLLKQAKATAETPPRKDLALLPPKKEEPTKPLFKLLG 681  
DB 1059 KDLNEILLVHYPSQKLLKARMLRSNNK-----PKEKSVLILPPRAGTPKFLNAA 1113  
QY 682 GTGKAS-----LARKLLKLRQAAQKKNSEEGEGEKENEDKQKEDKQKE 729  
DB 1114 MTGKMGKGGKAGKGLAHLRLARLKEALAEAAKQQL---VEQAKSODVKGEGSAAPD 1170  
QY 730 NEDKGKE-NEDKDKGREPEKP-----LDRPECTASPAVEEPPHSVTRVLPR 777

DB 1171 QHTHPKEAATDPAPRTTPPEPGSPSPSPASLGRPEGEGCA-EPEEHSVR----- 1223  
QY 778 GTSRQLIISMAPSARGGSEVLTIKVKAKQ 809  
DB 1224 -----ICMSFGPEPGEQILSVKMPEREE 1247  
RESULT 10  
ABG27471  
ID ABG27471 standard; protein; 652 AA.  
XX AC ABC27471;  
XX DT 18-FEB-2002 (first entry)  
XX DE Novel human diagnostic protein #27462.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS Homo sapiens.  
XX PN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US0008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Drmanac RT, Liu C, Tang YT;  
XX WPI: 2001-639362/73.  
XX N-PSDB: AAS91658.  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
PS Claim 20; SEQ ID NO 57830; 103pp; English.  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)  
XX  
SQ Sequence 652 AA;

Query Match 42.1%; Score 1781.5; DB 4; Length 652;  
Best Local Similarity 55.3%; Pred. No. 1.1e-142;  
Matches 343; Conservative 87; Mismatches 143; Indels 47; Gaps 7;



QY 700 AOKKENS 706  
 Db 956 MRRKQS 962  
 RESULT 12  
 ABG05466  
 ID ABG05466 standard; protein; 644 AA.  
 XX  
 AC ABG05466;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #5457.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US008631.  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Drmanac RT, Liu C, Tang YT;  
 PI  
 XX WPI: 2001-639362/73.  
 DR N-PSDB; AAS69653.  
 DR  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensic, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 PT  
 XX Claim 20; SEQ ID NO 35825; 103pp; English.  
 PS  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensic, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 644 AA;  
 Query Match 25.7%; Score 1088; DB 4; Length 644;  
 Best Local Similarity 100.0%; Pred. No. 2e-83;  
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 594 SLLAAGGNNRTANVAHGFANLLTLDKKTQLQELVHYPSERILMKARVLLKQAKTA 653

Db 429 SLLAAGGNNRTANVAHGFANLLTLDKKTQLQELVHYPSERILMKARVLLKQAKTA 488  
 QY 654 EATPPRKDLALLFPKKEETPKLFTLLGGTGKASLARLLKLRQAAOKKENSEGEEEG 713  
 Db 489 EATPPRKDLALLFPKKEETPKLFTLLGGTGKASLARLLKLRQAAOKKENSEGEEEG 548  
 QY 714 KENEDKQKENEKDKENEDKDKGREPEEKPLDRPECTASPIAVEEPEHVSVRT 773  
 Db 549 KENEDKQKENEKDKENEDKDKGREPEEKPLDRPECTASPIAVEEPEHVSVRT 608  
 QY 774 VLPRTGTSRQSLIISMAPSAEGGEEVLTIEVKEKAKQ 809  
 Db 609 VLPRTGTSRQSLIISMAPSAEGGEEVLTIEVKEKAKQ 644  
 RESULT 13  
 ADZ88294  
 ID ADZ88294 standard; protein; 821 AA.  
 XX  
 AC ADZ88294;  
 XX  
 DT 14-JUL-2005 (first entry)  
 XX  
 DE Human CNG channel proteins consensus sequence.  
 XX  
 KW antiasthmatic; Antiallergic; Hypotensive; Gastrointestinal-Gen.;  
 KW Antiparkinsonian; Antidepressant; Neuroleptic; Cardiant; GPCR-Antagonist;  
 KW Unspecified GPCR; cyclic nucleotide-gated channel A2; CNGA2; CNG;  
 KW mutation; divalent cation-mediated blockage; cation flux;  
 KW G-protein coupled receptor; asthma; hypertension; allergic reaction;  
 KW Parkinsons disease; depression; schizophrenia; heart disease.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 1..821 /label= Xaa = any amino acid  
 FT  
 PN WO2005041758-A2.  
 XX  
 PD 12-MAY-2005.  
 XX  
 PF 01-NOV-2004; 2004WO-US036022.  
 XX  
 PR 30-OCT-2003; 2003US-0515442P.  
 PR 20-JUL-2004; 2004US-0589012P.  
 XX  
 XX (ATTO-) ATTO BIOSCIENCE INC.  
 PA  
 XX Yao Y, Cao L, Lu J, Llorente I;  
 XX WPI: 2005-366496/37.  
 DR  
 XX New isolated polynucleotide encoding a mutant cyclic nucleotide-gated  
 PT (CNG) channel having a mutation that decreases divalent cation-mediated  
 PT blockage of cation flux, and makes the channel more sensitive to CAMP.  
 XX  
 PS Disclosure; SEQ ID NO 19; 156pp; English.  
 XX  
 CC This sequence represents a consensus sequence of the human cyclic  
 CC nucleotide-gated (CNG) channel proteins, CNGA1 - CNGA3 and CNGB1 - CNGB3.  
 CC The rat CNG channel sequence, CNGA2, may be mutated to form the isolated  
 CC polynucleotide of the invention which comprising a sequence encoding a  
 CC mutant CNG channel comprising at least one mutation that decreases  
 CC divalent cation-mediated blockage of cation flux, and at least one  
 CC mutation that makes the channel more sensitive to CAMP than a channel  
 CC that does not comprise the mutation. The invention includes methods for:  
 CC (a) detecting changes in intracellular levels of CAMP, comprising  
 CC expressing in a cell the polynucleotide cited above, and measuring  
 CC activity of the channel, where activity of the channel is indicative of  
 CC changes in intracellular CAMP; (b) detecting activity of a GPCR.



Db	224	SFSPIWFVDCDDPNLTSVLRPEDTENNYYLRQOLKCTWFMELDQPLPMGQNGKIEVTS	283
QY	303	-----	302
Db	284	EBEEMAGDIKYLNHCKWKKEELISGKLEDEGIEKENLATLEKIRKTSQGHNLNVMVC	343
QY	303	-----	302
Db	344	LGSGETTKPRNSGSTRRKESKQLKAGSQREWLVQIQTIGILMSEKAELWTALYYTQHA	403
QY	303	-----DVASITPF-----	310
Db	404	ARQPEGESKDLASCLQYSRWVQDNEGLSLNREQEERLLELEQKAEWEEQVEVHRQTL	463
QY	311	-----DICVLPFGFNPFRANRMLKYTSFFP-----	336
Db	464	TWQNDLTTISHAVFONGELKEQLAKLQTGFPMKLNNENWEITSTPQSEQHKKPESEKLG	523
QY	337	-----FNH	339
Db	524	GLDSEEEALPLMPSIPEELESREAMVAFNLNSAVASAEEOQAWPRQLKEQKGAMEKLQSH	583
QY	340	HLSEMDKAYIYR-----	352
Db	584	FMELMQEKVDLKEQMEKLERRCIQLSGETDVIHHTVPELEVSADVAPEGIHQQFPGQ	643
QY	353	-----	352
Db	644	GDGEERAGAAGAAGAGGIPERGIVIGNDSSTHVMPEDLPGQDVEDSDIDDP	703
QY	353	---VIRTTGYLLFILHINACVYVMASNYEGIGTTRWVYDGEENE	393
Db	704	DPVIRTTGYLLFILHINACVYVMASNYEGIGTTRWVYDGEENE	747

RESULT 15

ABR39396

ID ABR39396 standard; protein; 237 AA.

XX AC ABR39396;

XX AC

DT 26-JUN-2003 (first entry)

XX Amino acid sequence of human beta1b.

DE

XX

KW OCN1; OCN2; beta1b; CNG; human;

KW olfactory cyclic nucleotide gated channel subunit.

XX

OS Homo sapiens.

XX

XX WO2003004611-A2.

PN

PD 16-JAN-2003.

XX

PF 08-JUL-2002; 2002WO-US021184.

XX

PR 06-JUL-2001; 2001US-0303140P.

PR 10-DEC-2001; 2001US-0337154P.

XX

XX (SENO-) SENOMYX INC.

PA

XX Zoller MT, Xu H, Staszewski L, Moyer B, Pronin A, Adler JE;

PI Servant G, Callamaras N;

PI

XX WPI; 2003-229406/22.

DR

DR N-PSDB; ABZ80565.

XX

XX Novel isolated nucleotide sequences encoding human OCN1, OCN2, beta 1b

PT olfactory cyclic nucleotide gated (CNG) channel subunits, useful for

PT identifying the CNG channel activators useful for enhancing smell.

XX

PS Disclosure; Page 86; 97pp; English.

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2005, 19:47:49 ; Search time 43 Seconds  
(without alignments)  
1810.218 Million cell updates/sec

Title: US-09-855-828-1  
Perfect score: 4234  
Sequence: 1 MFKSLTKVKKVPDIGNENN.....PSABGGSEVLTIEVKEKAKQ 809

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1952	46.1	909	2 S32538	cGMP-gated cation
2	1095	25.9	800	2 T19627	hypothetical prote
3	834.5	19.7	695	2 S74179	cyclic nucleotide
4	829	19.6	663	2 S11521	CAMP-gated channel
5	821.5	19.4	735	2 I50630	alpha subunit of c
6	820	19.4	732	2 S35691	cyclic nucleotide
7	819.5	19.4	688	2 B42161	cGMP-gated cation
8	818.5	19.3	690	2 A42161	cGMP-gated cation
9	818	19.3	685	1 A44842	cGMP-gated ion cha
10	807.5	19.1	682	1 JH0560	cyclic nucleotide
11	804	19.0	664	2 T11517	cyclic nucleotide
12	801.5	18.9	691	2 J06509	rod cyclic nucleot
13	797	18.8	690	1 S27103	cGMP-gated ion cha
14	794	18.8	706	2 A55251	cyclic nucleotide
15	771	18.2	645	2 I50680	alpha subunit of r
16	770	18.2	772	2 S28232	hypothetical prote
17	718	17.0	665	2 S52072	DmCNGC protein - f
18	640	15.1	575	2 I59327	olfactory cyclic n
19	555	13.1	673	2 T20936	hypothetical prote
20	530	12.5	611	2 T20935	hypothetical prote
21	484	11.4	261	2 I78559	cyclic nucleotide
22	476	11.2	261	2 I78560	cyclic nucleotide
23	441.5	10.4	644	2 T33125	hypothetical prote
24	366.5	8.7	767	2 T21969	hypothetical prote
25	349	8.2	828	2 T52046	potassium channel
26	349	8.2	962	2 I53197	potassium channel
27	336	7.9	989	2 I48912	potassium channel
28	331.5	7.8	807	2 T12177	potassium channel
29	323.5	7.6	845	2 T07052	probable potassium

30	321	7.6	688	2 S55349	potassium channel
31	318	7.5	787	2 S68699	potassium channel
32	317.5	7.5	1159	2 I38465	probable potassium
33	314	7.4	883	2 T07651	potassium channel
34	313.5	7.4	662	2 T04461	potassium channel
35	311	7.3	887	2 T03939	potassium channel
36	302.5	7.1	838	2 S23606	potassium channel
37	302.5	7.1	857	2 S62694	potassium channel
38	296	7.0	934	2 T42394	potassium channel
39	292	6.9	1174	2 A40853	potassium channel
40	287	6.8	716	2 T51354	cyclic nucleotide
41	284	6.7	706	2 T51432	cyclic nucleotide
42	282	6.7	880	2 F85381	potassium channel
43	282	6.7	916	2 T05360	probable potassium
44	280	6.6	702	2 T04424	probable calmoduli
45	278.5	6.6	718	2 G84902	hypothetical prote

ALIGNMENTS

RESULT 1

S32538  
cGMP-gated cation channel 2, rod - human  
C;Species: Homo sapiens (man)  
C;Date: 22-Nov-1993 #sequence\_revision 17-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: S32538  
R;Chen, T.Y.; Peng, Y.W.; Dhallan, R.S.; Ahmed, B.; Reed, R.R.; Yau, K.W.  
Nature 362, 764-767, 1993  
A;Title: A new subunit of the cyclic nucleotide-gated cation channel in retinal rods.  
A;Reference number: S32538; MUID:93226050; PMID:7682292  
A;Accession: S32538  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-909 <CHE>  
A;Cross-references: UNIPROT:O43636; UNIPARC:UPI000017C0CA  
C;Keywords: cGMP binding  
F;620-742/Domain: CAMP receptor protein cyclic nucleotide-binding domain homology <CAP>

Query Match 46.1%; Score 1952; DB 2; Length 909;  
Best Local Similarity 46.4%; Pred. No. 2.3e-113;  
Matches 405; Conservative 121; Mismatches 237; Indels 110; Gaps 18;

Qy	14	IGENNENEQSRNREGSHPSNOSQQTAGE--ENKGE---ENKSLTKTSPTVSEEPHTN	68
Db	67	VGEAKKEAEKAEAEVAEEAEKEPQDMAETKEEPEAEAEAAASSGVPATKQHPEVQ	126
Qy	69	IQDK-----LSKXSSGDLTTNPPQNAAEPTGTVPQEKMDPGKGN-----	112
Db	127	VEDTDADSCPLMAEENPPS--TVLPSPSPAKSDTLIVPSSASGTHRKKLPSDEAEELK	184
Qy	113	--SPQNKP-----PAAP-----VINEYADAQLHNLVKRMORTALY	146
Db	185	ALSPAESPVVAMSDPTTPKDTQGDRAASTASTNSAIIND----RLQELVKLPKERTKV	240
Qy	147	KKKLVGDLG-----SPEASQTAKPATVPPVKESDDKPT--EHYRLLWFKVKMPLTEY	200
Db	241	KEKLIDPDVTSDBESPSPAKKAPAPDTPKPAEAPVEEHCYDMCLCKFKHRPMKKY	300
Qy	201	LKRKLPNSIDSVTDRLYLALJLLVLAYNMNCWFIPLRLVFPYOTADNLIHYMLADIIC	260
Db	301	-----QFPQSIDPLTNLMYLVLFVVMWNNWNLIPVRWAFVPTPDNIHMLMDYLC	356
Qy	261	DIIVLYDMLFIQPLQFVRGDIIVDSNELRKHYRTSTKFLQDVASIPEDICVLFPGFN	320
Db	357	DLIVFLDITVFQFELQFVRGDIITDKDMNNYLSRPFKMDLLSLPLDFLYLKGVN	416
Qy	321	PMFRANMLKYTFPFENHLESIMDKAYIYRVIRTTGYLLFILHINACVYVWASNEGI	380
Db	417	PLLRPLRCLXYMAPFEENSLESILSKAYVYRVIRTTAYLLYLSLHNSCLYVWASVQGL	476
Qy	381	GTTFRWYDGGSENLRCYYWAVRTLITIGGLPEPOTLFEIVFOLLNFFSGVVFSSLIQO	440

```
Db 477 GETHVYDGVNSYRCYYFAVKTLITIGGLPDPKTLFEIVFQLLNYFTGVFAFSWMICQ 536
Qy 441 MRDVI GAATANQNYFRACMDTTIAYNNYSIPKLQKRVRTYVEYTWDSQRMLEDSDLLK 500
Db 537 MRDVVGAATAGQTYYSCHDSIVKYNFYKIPKSVQNRVKTYEYTWHSQGLDSELMV 596
Qy 501 TLPTTVQLALADVNPISIISKVDLPKGCDDQMIYMLLKLKSVLYLPDGFVCKKGIGKE 560
Db 597 QLPDRKMLDLAIDVNNIVISKVALFOGCDQRMIFDMLKELRSVYVLPNDYVCKKGIGRE 656
Qy 561 MYIIKHGEVQVLGGDGTGVLVTLKAGSVFGEISLLAAGGNRRRTANVAHGFANLLTLD 620
Db 657 MYIIQAGQVQVLGGDGTGVLVTLKAGSVFGEISLLAAGGNRRRTANVAHGFANLLTLD 716
Qy 621 KKTLOEILVHYDPDSERILMKKARVLLKOKATAEATPPRKDLALLFPKPKFTKLL 680
Db 717 KKDLEILVHYDPESOKLLRKARRMLRSNNK-----PKEKSVLLLPAPGATPKLFNAAL 771
Qy 681 GGTGKAS-----LALLKLRQAAQKKNSEGGEEBEGKEDKOKEDKOK 728
Db 772 AMTGRMGKGKAGKGLAHLARLRLKELAALEAAAKQOEL---VEQAKSSQDVKGEESAAP 828
Qy 729 ENEDKGE-NEQDKGREPEKP-----LDREPTASPTAVEEPEHSHSVRTVLP 776
Db 829 DQHTPKAEATDPPAPRTPEPPGPPSPPPASLGRPEGBEIPA-EPEHSHSV-----882
Qy 777 RGTSRQSLIISMAPSAEGGEVLTIEVKEKAG 809
Db 883 -----ICMSPGPEGEQILSVKMPERE 906

RESULT 2
T19627
hypothetical protein F36F2.5 - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C.Accession: T19627; T21863
R:Kershaw, J.
A:Reference number: Z19153
A:Accession: T19627
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-800 <W1>
A:Cross-references: UNIPROT:O622237; UNIPARC:UPI0000083D03; EMBL:Z93778; PIDN:CAB07847.1;
A:Experimental source: clone C31H5
R:Cottage, A.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19479
A:Accession: T21863
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-800 <W12>
A:Cross-references: UNIPARC:UPI0000083D03; EMBL:Z81532; PIDN:CAB04328.1; GSPDB:GN000019;
A:Experimental source: clone F36F2
C:Genetics:
A:Gene: CESP:F36F2.5
A:Map position: 1
A:Introns: 27/3; 51/2; 142/2; 191/3; 287/3; 333/3; 355/2; 405/1; 433/2; 484/1; 508/1; 56

Query Match 25.9%; Score 1095; DB 2; Length 800;
Best Local Similarity 31.9%; Pred. No. 3.4e-60;
Matches 258; Conservative 153; Mismatches 301; Indels 98; Gaps 22;

Qy 2 FKSLLTKVNVKPIGENENEQSSRRNEEG---SHPSNQSQQTAAQENKGEKSLTKST 58
Db 21 FSVDRQKASKPT-QLSEKWKSPRSDSFLLDPANAS-----KEPSASTRL 68
Qy 59 PV-TSEEPHTNQKLSKKNSSGDLTTNP-----DPQNAAEPTGTVPQEKMDPGKEGN 112
Db 69 PYPPTRPPEVVIQI-----DEVESILGLIDETDDHLDG-----RLDPASSFDA 113
Qy 113 SPQNPAPPAAPVINEYADAQ-----LHNVKRMQRQTALYKKKLVLGEGDLSSEASPTQ 164
```

```
Db 114 NLSATRASSIIEDVRSQISPIMRERLHSAKEVHRRTSAVREDLIR---ETPEDTVSM 170
Qy 165 AKPTAVPPVKSGDDKPTHEHYRLWFKVK-KMPLTEYLKR---IKLPNSIDSYTDRLYLL 220
Db 171 A--SNVP--KONEHRPS--LMSLIGLQNRSESPVTDVKNCFGSLKGTFFHY-GRFYMT 223
Qy 221 WLLVLTAYNNWCWPIRLRVPFYQTADNIHWLIADIICDIIYLYDLFIQPRIQFVRG 280
Db 224 WLSLVTLCFLFNACFIPLRSSYPYQTADNMWYFVDYSCDLVYVIMLLIKPRLRFTRG 283
Qy 281 GDIIVDSNELKRYSTKTFQDVASIIPDICYLFFGFNPMFRANRMKYTSPEFNNH 340
Db 284 GIQVIYKDTORHYLMNTRTFKLDILSLPTDLMVFFFGKMPITWIRNRLVKNPSFLLFDM 343
Qy 341 LESINDKAYIVRVTRTTCYLLFIHLINACVYMASNYEGIG-----TTTRWVVDG 389
Db 344 LQNSFANPYAIRIAKTLSTMIYIHCNSCVYKUSALQAFQAIAYLENGKMYLNKMWYNN 403
Qy 390 EGNELRCVYMAVRLTITIGGLPEPOTLFEIVFOLLNFFSGVFPFSSLIQMRDVI GAAT 444
Db 404 QGNAYIRCFYPTAAVATSTGNNPAPTNVIEIYMTCSMMMGVFPFALLLGOIRDIVSNAN 463
Qy 450 ANQNYFRACMDDTIAYMNNYSIPKLQKRVRTYVEYTWDSORMLDESOLLKTLPTTVQLA 509
Db 464 RNREEFORKMDLALGECKKLGKMETTNRVRDWFYTWQOQKTLDEKKLIEKLPILQTD 523
Qy 510 LAIDVNFISIISKVDLPKGCDDQMIYMLLKLKSVLYLPDGFVCKKGIGKEMYIHKHGFV 569
Db 524 LALSVMHTYTLTSKVQLFQCDRALLRDLVLKLRPVIFLPGDMICLKGVDGKEMYIINOGLI 583
Qy 570 QVLGGPDGTVLTLKAGSVFGEISLLAAGGNRRRTANVAHGFANLLTLDKKTIQEILV 629
Db 584 QVVGSDHNEKIFAEALAQAVFGEISLLAAGNNRRTASIRAKGYCTLFVLAKEDLNDVIR 643
Qy 630 HYPDSERILMKARVLLKQAKTAEATPPRKDLALL-----FPPKEETPKFCTLGSGT 683
Db 644 YYPQATILRRKAAAMLNKKSKDEKTEKIKAQAELEDRCCKINPR-QVPKLITLIANMTE 702
Qy 684 -----GKASLARLLKLKREAAQKK-----ENSEGEEBEGKEDKOKENED 724
Db 703 MNENKGVQELKKVIEEETEKRRRSIYYPWSTLQRDDDEBEWNEDEEDLSOVGEDFDLP 762
Qy 726 -KQENKGDGKENEKDKGREPEKPLDRP 754
Db 763 TNHSDDED---PMEDVDLAPVHDDDDWQDP 789

RESULT 3
S74179
cyclic nucleotide-gated channel protein - human
C.Species: Homo sapiens (man)
C.Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 16-Jul-1999
C.Accession: S74179
R:Yu, W.P.; Grunwald, M.E.; Yau, K.W.
FEBS Lett. 393, 211-215, 1996
A>Title: Molecular cloning, functional expression and chromosomal localization of a human
A:Reference number: S74179; MUID:96409310; PMID:8814292
A:Accession: S74179
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-695 <YUW>
A:Cross-references: UNIPARC:UPI00000177899
A:Experimental source: retina
C:Genetics:
A:Map position: 2
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-gated
C:Keywords: ion channel; ion transport; membrane protein
F:482-607/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology (CA)
Query Match 19.7%; Score 834.5; DB 2; Length 695;
Best Local Similarity 29.7%; Pred. No. 4.3e-44;
Matches 223; Conservative 149; Mismatches 271; Indels 107; Gaps 22.
```





```

Db      3  LSKMTNIINTWTSFVNIPNVIIPVAIEKEIRRMWENGACSFSDDDN-----GSLSESENE 57
Qy      106 PG-----KEGPNSPQNKPPAAPV---INEYADAQLHNLVKRMQRORTALYKXKLVLEGD 154
Db      58  DSFFRSNSYKRRGPGSQREQHLPGTMAIFNVNNSNKKDQEPKPKKKKK-----EKSKADD 113
Qy      155 LSPSPASPTAKTAVPPVKESDDKPTHEHYRLMLFPKVKQKMPLETVLKRKIKLPNSIDSYT 214
Db      114  KNEKNKDKPEKKKKE-----KKEKKKKKE-----KKEKKKEKEKEVVVVDPSGNTYY 162
Qy      215  DRLYLMLLLVTL---AYNNCNWFILPLRVPYQTADNIHYWLIADIICDIYLYDMFLTO 272
Db      163 N-----WLFCTILFVWYNWT---MIARACFDELQSDYLEYWLIFDVSNVVIADLM-FVR 214
Qy      273  PRLQFVRGGDIIVDSNELKRYRTSTKFOLDVASIIPFDICYLFFGFN-PMFRANRMLKY 331
Db      215  TRTGYLEQLLVKORMKLEKYKANLQFKLDVLSVPTDLLYIKFGWNYPEIRLNRLAI 274
Qy      332  TSFFFNHLESIMDKAYIRVIRTTGYLLFILHINACVYVWASNYEGIGTTRVY-----387
Db      275  SRMFEPFOTETRTNYPNIFRISNLVMYVIIHWNACVYSISKAIKFGNDTWVYPDVN 334
Qy      388  DGE-----GNEYLECYWAVRTLTITIGLPEPQTLFEIVFQLNFFSGVVFSSLIQOMRD 443
Db      335  DPEFGRLARKYVYSLYWSLTLTITIGLPEPQTLFEIVFQLNFFSGVVFSSLIQOMRD 394
Qy      444  VTGAATANQNYFRACMDTIAVNNYSIPKLQKRVRTWYETWDSQRMDESLLKTLTP 503
Db      395  MTSNNMAARAEFQSRVDAIKQYNNFRNVSKMEKRVIKWFDYLTWTKTVDREVLRYLP 454
Qy      504  TTVQLALAIIVNFSIISKVDFKGCOTMIYDMLLKSLVLYLPQGVFCKKGIGKEMYI 563
Db      455  DKLRABIAINVLHDLTKVRIFADCEAGLLVELVLKLPQVYSPGYICKKGDIGREMYI 514
Qy      564  IKHGEVOVLGGPGTKVLYTLKAGSVGEISLLAAG---GNRRTANVAHGFANLLTLD 620
Db      515  IKEGKLAV-ADDGITQFVLSGVSFGEISILNKGKAGNRRRTANIKSIGYSDLFCLIS 573
Qy      621  KKTQILVHYDPDSERILMKKARVLL 646
Db      574  KODLMEVLYEYPAKTMLEKGRQIL 599

RESULT 8
A42161
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C:Accession: A42161
R:Pittler, S.J.; Lee, A.K.; Alther, M.R.; Howard, T.A.; Seidin, M.F.; Hurwitz, R.L.; W
J. Biol. Chem. 267, 6257-6262, 1992
A:Title: Primary structure and chromosomal localization of human and mouse rod photorece
A:Reference number: A42161; MUID:92210603; PMID:1372902
A:Accession: A42161
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-690 <PT>
A:Cross-references: UNIPROT:P29973; UNIPARC:UPI0000163B2A; CB:MB4741; NID:q180461; PIDN:
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-
C:Keywords: cAMP binding; transmembrane protein
F;479-603/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP>

Query Match 19.3%; Score 818.5; DB 2; Length 690;
Best Local Similarity 30.08; Pred. No. 4.2e-43;
Matches 221; Conservative 128; Mismatches 262; Indels 125; Gaps 22;

Qy      5  LTKVNVKPKIGENNENQSSRRNEEGSHPS-----NQSQQTAAQENKGE-----KSL 53
Db      17  VTPNVPV-----DIEKIRRMWENGACSFSDDDSDAYSITSESENENPHARGSFYSKSL 71
Qy      54  KTKSTPTVTSSEPHNTIQDKLSKNSSGDLTTNPPQNAEPTGTVPQKEMDPKSGPNS 113

```

```

Db      72  R-KGQPSQREQVLPGLAIPIFNVNNS-----NKD-----QEPEKKKKKKKKSSDDKNE 121
Qy      114  POKPAPAPVINEYADAQLHNLVKRMQRORTALYKXKLVLEGDLSSEASPTAKTAVPPV 173
Db      122  NKNDPE-----KKKKKKKKKK-----E 141
Qy      174  KESDDKPTHEHYRLMLFPKVKQKMPLETVLKRKIKLPNSIDSYTDLRLLMLLVTL---AYNW 231
Db      142  EKSKDKKHH-----KKEVVV-----IDP-SGNTYYNWLFCITLPMVYWN 180
Qy      232  NCWFILRLVFPYQTADNIHYWLIADIICDIYLYDMFLQIPRLQFVRGGDIIVDSNELR 291
Db      181  T---MYIARACFDELQSDYLEYWLILDYVSDIVYLIDM-FVTRTGYLEOGLLVKEELKI 237
Qy      292  KHYRTSTKFPQDVASIIIPFDICYLFFGFN-PMFRANRMLKYTSFPEFNHLESIMDKAYI 350
Db      238  NKYKSNLQFKLDVLSVPTDLLYIKFGWNYPEIRLNRLRFRSMFEFFORTETRTNYPNI 297
Qy      351  YRVIRTTGYLLFILHINACVYVWASNYEGIGTTRVY-----DGE-----GNEYLECYWAV 402
Db      298  FRISNLVMYVIIHWNACVYSISKAIKFGNDTWVYPDINDEPFGRLARKYVYSLYWS 357
Qy      403  RTLTITIGLPEPQTLFEIVFQLNFFSGVVFSSLIQOMRDVIGAATANQNYFRACMDT 462
Db      358  LTLTIGETPPVRDSEYVYVVDLIGVLIFATIVGNIGSMISNNMAARAEFOARIDA 417
Qy      463  IAYMNNYSIPKLQKRVRTWYETWDSQRMDESLLKTLPTTVQLALAIIVNFSIISKV 522
Db      418  KQYMHFRNVSKMEKRVIKWFDYLTWTKTVDREVLRYLP 477
Qy      523  DLFPKGCOTMIYDMLLKSLVLYLPQGVFCKKGIGKEMYI IKHGEVOVLGGPGTKVLY 582
Db      478  RIFADCEAGLLVELVLKLPQVYSPGYICKKGDIGREMYI IKEGKLAV-ADDGITQFV 536
Qy      583  TLKAGSVGEISLLAAG---GNRRTANVAHGFANLLTLDKKTQILVHYDPDSERILM 639
Db      537  VLSGVSFGEISILNKGKAGNRRRTANIKSIGYSDLFCLSKDDLMLEALTEYDPAKTM 596
Qy      640  KKARVLLKOKA---KTAETPPRKDALLPFPKKEETPKLFKTLGCGTGKASLARLL--- 692
Db      597  EKGKQILMKDGLDLNANAGSDPKDLBEKVTMEGSDVLLQ-----RPARILAEY 648
Qy      693  -----KLKREQAQKK 703
Db      649  ESMQKQLKQRLTKVEK 664

RESULT 9
A44842
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A44842
R:Dhallan, R.S.; Macke, J.P.; Eddy, R.L.; Shows, T.B.; Reed, R.R.; Yau, K.W.; Nachans, J
J. Neurosci. 12, 3248-3256, 1992
A:Title: Human rod photoreceptor cGMP-gated channel: amino acid sequence, gene structure
A:Reference number: A44842; MUID:92356211; PMID:1379636
A:Accession: A44842
A:Molecule type: mRNA
A:Residues: 1-686 <DHA>
A:Cross-references: UNIPARC:UPI000008787; GB:942457; NID:9252853; PIDN:AA822778.1; PID:
A:Experimental source: retina
A:Note: sequence extracted from NCBI backbone (NCBIN:110250, NCBIP:110251)
A:Note: intron positions were determined from genomic sequence
C:Genetics:
A:Gene: GDB:CNCG1; CNCG
A:Cross-references: GDB:127557; OMIM:123825
A:Map position: 4p14-q13
A:Introns: 36/2; 75/2; 95/2; 110/2; 146/2; 182/2; 218/2
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-
C:Keywords: cAMP binding; eye; ion channel; ion transport; retina; transmembrane protein
F;475-599/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP>

```

Query Match 19.3%; Score 818; DB 1; Length 686;  
Best Local Similarity 29.9%; Pred. No. 4.4e-43;  
Matches 226; Conservative 129; Mismatches 266; Indels 136; Gaps 24;

QY 5 LTKVNVKPIGENNEQSSRRNEGSHPS-----NQSQTTAQRNKGBE-----KSL 53  
DB 13 VTMPIVIVP-----DIEKEIRMERGACSSFSDSDSASTSESENPHARGSFYSKSL 67  
QY 54 KTKSPVTSSEPHNTIQDKLKKNSGDLTTPDQNAAPETGTVPQEKMEPGKGGNS 113  
DB 68 R-KGSPSREQVLPQALFVNNSS-----NKD-----QEPKPKKKKKKKSKDNKKE 117  
QY 114 POKKPPAAPVINEYADAQLHNLVKRMQRORTALYKKLVVEGDLSSPEASQTAKTAVPPV 173  
DB 118 NKNDPE-----KKKKKKKKKKK-----E 137  
QY 174 KESDDKPTHEYVRLWFKVKMPLTEYLKRIKLPSIDSYTDRLYLMLLVTL--AYNW 231  
DB 138 EKSKDKKEE-----KKEVV-----IDP-SGNTYNNLFCITLPMVMYMW 176  
QY 232 NCWFPLRLVFPYQTADNIHYWLIADIICDIIVLYDMLFIQRLQFVRGGDIIVDSNELR 291  
DB 177 T--WVIARACFDELQSDYLEWLIIDYSDIVLIDM-FVTRTGYLEGQLLVKSELKI 233  
QY 292 KHYRTSTKQOLDVASIIPDICYLFFGFN-PMFRANRMLKYTSFFEFNHLESIMDKAYI 350  
DB 234 NKYKSNLQFKLDVLSIPTDLLYFKLGWNPYPIRLNRLRFRSMFEFFORTETRTNYPNI 293  
QY 351 YRVIRTTGYLLFIHINACVYVWASNYEGIGTRWVY-----DGE-----GNEYLRCYVAV 402  
DB 294 FRISNLVYIVIIHWNACVFYSISKAIGFGNDTWYDINDPFRGLARKYVYSLYMS 353  
QY 403 RLTIITIGLPEPQTLFEIVFOLLNFFSGVVFSSLIGOMRDVIGAATANONFRACMDT 462  
DB 354 LTLTIGETPPVRDSEYVVFVDFELGLVIFATIVNGISMISNMNARAEPQARIDAI 413  
QY 463 IAYMNYSIPLKVKRVRTWBYTWSQRMDESLLKTLPTTVQALAIIDVNFISIISKV 522  
DB 414 KQYMFHNVSKDMERVKVFKFDYLTNKKTVDEKVLKYLDPKRAEIAINVHLDLTKKV 473  
QY 523 DLFKGCDDQMIYDMLRLKSVLYLPGDFVCKKEIGKEMYIIKHGEVOVLGGPDGTVLV 582  
DB 474 RIFADCEAGLVVELVLPQVYSPGDYICKGDIKREMYIIKEGQAVV-ADGGVTOFA 532  
QY 583 TLKAGSVFGEISLLAAGG---GNRRATANVAHGFANLLTLDKKTQLQELVHPDSEIIM 639  
DB 533 VLSDSYFGEISILNIKSGKAGNRTANIKSIGSDLFCLSKDDLMEALTEYDPAKTMLE 592  
QY 640 KQARVLLKQKA-----KTABATPPKDLALLFPKKEETPKLFTKLGGTKASLARLL--- 692  
DB 593 EKGQILMKDGLLDLNIANAGSDPKLEBKVTRMEGSDVLLQT-----RFAILAEY 644  
QY 693 -----KLKREQAQK-----ENSEG-GESEG 713  
DB 645 ESMQOKLQRLTKVFKPLKPLIDTFSSIEGPGAESG 681

RESULT 10  
JH0560  
Cyclic nucleotide-gated channel - channel catfish  
C:Species: Ictalurus punctatus (channel catfish)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C:Accession: JH0560  
R:Goulding, B.H.; Ngai, J.; Kramer, R.H.; Colicos, S.; Axel, R.; Siegelbaum, S.A.; Chesel  
Neuron 8 45-58, 1992  
A:Title: Molecular cloning and single-channel properties of the cyclic nucleotide-gated  
A:Reference number: JH0560; PMID:92110008; PMID:1370374  
A:Accession: JH0560  
A:Status: nucleic acid  
A:Molecule type: mRNA  
A:Residues: 1-682 <GOU>  
A:Cross-references: UNIPROT:P55934; UNIPARC:UPI0000127C25; GB:M83111

A:Experimental source: olfactory epithelium  
C:Comment: This cyclic nucleotide-gated channel is activated equally well by both cAMP and  
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide  
C:Keywords: cAMP binding; cGMP binding; ion channel; ion transport; olfaction; transmembr  
F:137-157/Domain: transmembrane #status predicted <TS1>  
F:173-193/Domain: transmembrane #status predicted <TS2>  
F:217-236/Domain: transmembrane #status predicted <TS3>  
F:241-261/Domain: transmembrane #status predicted <TS4>  
F:277-297/Domain: transmembrane #status predicted <TS5>  
F:319-337/Domain: transmembrane #status predicted <TS6>  
F:350-370/Domain: transmembrane #status predicted <TS6>  
F:447-571/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP>

Query Match 19.1%; Score 807.5; DB 1; Length 682;  
Best Local Similarity 29.0%; Pred. No. 2e-42;  
Matches 192; Conservative 137; Mismatches 256; Indels 77; Gaps 17;

QY 119 PAAPVINEYAD-----AQLHNLVKMRORTALYKKLVVEGDLSSPEA-----SPOTAKPT 168  
DB 55 PSAEMLEAFTQRRPLARLVNLVLSUREWA---HKSLSVETE-QRPDSFLERPRGPOAANDQ 110  
QY 169 AVPPVKESDDKPTHEYVRLWFKVKMPLTEYLKRIKLPSIDSYTDRLYLMLLVTLA 228  
DB 111 SAAPA-----DAPKTF-----KERWEGFVVSQSDDIYYWLFIALA 148  
QY 229 --YNNCWFILRLVFPYQTADNIHYWLIADIICDIIVLYDMLFIQRLQFVRGGDIIVD 286  
DB 149 SLYNW--IMLVARACFDQLQDENFFLWGLDVLCDVIVLDTCLIRLRTGYLEQLLVKLI 205  
QY 287 SNELKHYRTSTKQOLDVASIIPDICYLFFGFNPMFRANRMLKYTSFFEFNHLESIMD 446  
DB 206 LAKLRDNYIRTLOPKLDFLSILPTTELLFFVTGYVLPQRFNRLRFRSMFEFFDRTETRN 265  
QY 347 KAYIVRVIRTTGYLLFIHINACVYVWASNYEGIGTRWVYDGEKNE---YLRCYVAV 402  
DB 266 YNAPFRICNLILYLVIIHWNACIYALSKALGSSDTWVYSGONKTLSCFYVCYFWST 325  
QY 403 RLTIITIGLPEPQTLFEIVFOLLNFFSGVVFSSLIGOMRDVIGAATANONFRACMDT 462  
DB 326 LTLTIGEMPPPVKDEEYVVFVDFLVGLVIFATIVGNVGMIANMNATRAEFQTRIDAI 385  
QY 463 IAYMNYSIPLKVKRVRTWBYTWSQRMDESLLKTLPTTVQALAIIDVNFISIISKV 522  
DB 386 KHYMFHNRVNTLETRVTKVDFYLTNKKTVDEQVCLKNLPDKRAEIAINVHLDLTKKV 445  
QY 523 DLFKGCDDQMIYDMLRLKSVLYLPGDFVCKKEIGKEMYIIKHGEVOVLGGPDGTVLV 582  
DB 446 RIFQDCEAGLVVELVLPQVYSPGDYICKGDIKREMYIIKEGQAVV-ADGGVTOFA 504  
QY 583 TLKAGSVFGEISLLAAGG---GNRRATANVAHGFANLLTLDKKTQLQELVHPDSEIIM 639  
DB 505 LLTAGCGFGEISILNIQSGKMGNRRRTANIRSIGYSDLFCLSKDDLMEAAVEAPDAQKVL 564  
QY 640 KQARVLLKQKA-----AOKKENSEGGEEG---KENEDKOKENEDKOKENEDK 736  
DB 565 ERGREILRKQGLLDESVAAGGLVIDTEBKVERLDASL-----DILQTRFARLLGEFTST 619  
QY 684 GKASLARLLKLEQA-----AOKKENSEGGEEG---KENEDKOKENEDKOKENEDK 736  
DB 620 QRRLLKQRTALERQLCHTGLGLLSDNEAGEHAGVPTHADIAHQPEHTRTSAETNSE 679  
QY 737 NE 738  
DB 680 EE 681

RESULT 11  
S11517  
Cyclic nucleotide-activated channel protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S11517  
R:Dhallan, R.S.; Yau, K.W.; Schrader, K.A.; Reed, R.R.

Nature 347, 184-187, 1990  
A:Title: Primary structure and functional expression of a cyclic nucleotide-activated ch  
A:Reference number: S11517; MUID:903701115; PMID:1697649  
A:Accession: S11517  
A:Molecule type: mRNA  
A:Residues: 1-664 <DHA>  
A:Cross-references: UNIPROT:Q00195; UNIPARC:UPI0000000CF3; GB:X55519; NID:G56791; PIDN:C  
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-  
C:Keywords: transmembrane protein  
F:456-580/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP>  
Query Match 19.0%; Score 804; DB 2: Length 664;  
Best Local Similarity 29.8%; Pred. No. 3.1e-42; Indels 108; Gaps 22;  
Matches 219; Conservative 126; Mismatches 286; Indels 108; Gaps 22;  
QY 50 ESKLTKSTPVTSEEPHTNIQDKLSKNSGDLTTNPDPQNAAEPTGTVPQEQ---EMDP 106  
DB 4 EKSGVYKSSPANN---HNHPPPSIKANGKDDHRAGSRPQSVAAADDTSPELQRLAEMDT 60  
QY 107 GKEGPNSPQNKPPAAPVINEYADAQLH-----NLVKMRQRTALYKKLVGEGDLSSE 159  
DB 61 PRGRGGFQIRVLGVIRDMANKNFRPEEPRDPSFLERPR-----GPE 104  
QY 160 ASPTAKPTAVPPKESDDKPTHEYHLLVFKVKMPLTEYLKRIKLPNSIDSYTDRLYL 219  
DB 105 LQVTHHQGDDKGGKGGKTKKPFELF-----VLDPAGD-WYY 143  
QY 220 LWLLVTL--AYNWCWFIPLRLVFPYQTADNIHYWLIADIICDIIVYDMLFIQRLQF 277  
DB 144 RLFPVAMPVLYNW-CLLV-ARACFSDLQRYFVWLVLDYFSDTVYIAD-LIIRLTGF 200  
QY 278 VRGDIIVDSNEIRKHYRTSTKQFDVASIIPDICYLFFGP-NPMFRANMLKYTSFPE 336  
DB 201 LEQGLLVKDPKLRDNYHTLQKLDVASIIPDILYFAVGHSPVFRNLLHFARMPE 260  
QY 337 FNHLESIMDKAYIRVIRTTGYLLFILHINACVYVWASNYEGITRWY----DGE-- 390  
DB 261 FPORTETRTSYNIFRISNLVILYIIHWNACIYVYSKISGFGVDTVWYNPITDPEYG 320  
QY 391 --GNEYLCYVYAVRTLTITIGGLPEPOTLPEIYVQLNFFSGVFSSLIQMRDVIQAA 448  
DB 321 YLAREYIYCLYWSLTLTITIGETPPPVKDEYLFVIFDFLIGVLIIFATIVGVNGLSNM 380  
QY 449 TANQNYFRACMDDTIAYNNYSIPKLQKRVRTYETWDSQRLMDESLTLPTTVQL 508  
DB 381 NATRAEFOAKIDAVHYMQPRKYSKDEAKVIKWFYLTNKTVDREVLKLPKALRA 440  
QY 509 ALAIDVNFYSISKVDLPKCDTQMIYDMLRLKLSVLYLPGDFCKKGEIGKEMYYIKHGE 568  
DB 441 EIAINVHLSLTKKVRIFQDCEAGLLVELVLRPQVPSGDYICRKGDIKEMYYIKHGE 500  
QY 569 VQVLGGPDGKVLVTLKAGSVFGEISLLAAG---GNRRTANVVAHFANLLTLDDKKTQL 625  
DB 501 LAVV-ADGGVTOYALLSAGSFGFGEISILNKGSGWGNRTANIRSLGYSDLPCLSKDDL 559  
QY 626 EILVHPDSEIRILMKKARVLLKQAKTAE---ATPPRKDALLPKPEETPKFKTLGG 682  
DB 560 EAVTEYDPAKVLBERGRELKMGGLDENEVAASMEVDV-----QEKLEQETNN--D 611  
QY 683 TGAASLARLLKLKEQAQKENSEGEEBEGKEDKOKEN--EDKOKENEDKGENEDK 740  
DB 612 TLYTRFARLL-----AEYTAQOQKQIRITVLETRKQNH-----EDDYL 651  
QY 741 DKGREPEKPLDRPECTAS 759  
DB 652 SDG-----INTPEATA 663

RESULT 12

JC6509

rod cyclic nucleotide-gated cation channel protein alpha-chain - dog

C:Species: Canis lupus familiaris (dog)

C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004

C:Accession: JC6509

R:Veske, A.; Nilsson, S.E.G.; Gal, A.

Gene 202, 115-119, 1997

A:Title: Characterization of canine rod photoreceptor cAMP-gated cation channel alpha-sub

A:Reference number: JC6509; MUID:98087425; PMID:9427553

A:Accession: JC6509

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-691 &lt;VES&gt;

A:Cross-references: UNIPROT:Q28279; UNIPARC:UPI0000127C19; EMBL:X99914; NID:gl488571; PII

C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide

F:478-602/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology &lt;CAP&gt;

Query Match 18.9%; Score 801.5; DB 2: Length 691;

Best Local Similarity 29.2%; Pred. No. 4.7e-42;

Matches 220; Conservative 136; Mismatches 286; Indels 112; Gaps 23;

QY 21 EOSSRRNEGSHPSNOSOOTTAQEEENKGEKSLKSTPVTSEEPHTNIQDKLSKKNSSG 80

DB 24 EKEIRRMENGARSS-----FSDDDDGDDSDASMPFES---ENETPIA--RDSRRNSORR 72

QY 81 DUTTNPDQNAAEPTGTVPQEKEMDPGKEGPNSPQNKPPAAPVINEYADAQLHNLVKMR 140

DB 73 D-----PSQRE-----QVLPQALALFNVNSSNKEOEPEKK 104

QY 141 QRTALYKKLVGEGDLSSEASPTAKTAVPPVKSDDKPTHEYHLLVFKVKMPLTEY 200

DB 105 KKKK--EKSKGDKNENKDKSKKKKE---KEKEKNKKEK-----GKDKK---EEE 150

QY 201 LKRIKLPNSIDSYTDRLYLLWLLVTL--AYNWCWFIPLRLVFPYQTADNIHYWLIADI 258

DB 151 KKEWV---IDP-AGMYNWNLFCTILPVMYNT--MVIARACFDELOSDYLEWIIIFDY 204

QY 259 ICDIIVYDMLFIQRLQFVRGDIIVDSNEIRKHYRTSTKQFDVASIIPDICYLFFG 318

DB 205 LSDIVYLLDM-FVTRTGYLEQGLLVREBAKLEIKYKSNLQPKLQFSLVPTDLYFLKLG 263

QY 319 FN-PMFRANMLKYTSFPEFNHLESIMDKAYIRVIRTTGYLLFILHINACVYVWASNY 377

DB 264 WNYPEIRLNLRLIRISMFEFFQRTETRTNYPNIFRISNLVMIYVILIIHWNACVYFSIKA 323

QY 378 EGITGTRWY---DGE---GNEYLCYVYAVRTLTITIGGLPEPOTLPEIYVQLNFFS 429

DB 324 IGFNDTWYIPDVNDPEFORLARKYVYSLYWSLTLTITIGETPPPVDSYVFPVVDPLI 383

QY 430 GVFVSSLICQMRDVIQAAATANQNYFRACMDDTIAYNNYSIPKLQKRVRTYETWDS 489

DB 384 GVLIIFATIVGVNGLSNMNAARAEFOARIDAIKQYMHFRNVSKMEKRVIKWFDYLTN 443

QY 490 QRMLESDDLKTLPTTVQALAIIDVNFYSISKVDLPKCDTQMIYDMLRLKLSVLYLPGD 549

DB 444 KKTVDKELVKYLPDKLRAEIAINVHLDTLKKVRIFADCEAGLLVELVLRPQVPSGD 503

QY 550 FVCKKGEIGKEMYYIKHGEVOVLGGPDGKVLVTLKAGSVFGEISLLAAG---GNRRTA 606

DB 504 YICCKGDIGREMIYIKKGLAVV-ADGGITQFVVLSDGSYFGEISILNKGSKAGNRRTA 562

QY 607 NVVAHFANLLTLDDKKTQLQELVHPDSEIRILMKKARVLLKOKA---KTAETATPPRKDL 662

DB 563 NIKSIGYSDLPCLSKDDLMEALTEYDPAKTMLEKQKQILMKDGLLDINIANAGSDPKDL 622

QY 663 ALLFPKPEETPKFKTLGGTGKASLARLLKLKEQAQKENSEGEEBEGKEDKOKEN 722

DB 623 EEKVTMRGSDVLLQT-----RFAFIL-----ABYESMOOK 653

QY 723 NEDQKQENEDKGENEDK-----KGREPEEKPLD 752

DB 654 LKQRLTKVERFLKPIIDTFPSALEGTGDSRPLD 687

RESULT 13

S07103

cGMP-gated ion channel protein - bovine



Job time : 47 sec8

```
Qy 571 VLGGPDGKTVLTKAGSVGEISLLAAG--GNERTANVVAHFANLLTLDKKTQLQEI 627
Db 548 VW-AEDGITQVVLGDSYFGEISILNIGKSGNRTANIRISIGYSDLFCLSKDDLMEA 606
Qy 628 LVHYPDSERILMKAR-VLLKQK--AKTAEATPPRKDLALLPFPKPEETPKFKTLTGGT 683
Db 607 LTEYPEAKKALEEKGRIILMKDNLIDELAKAGADPKDI-----EEKVEHLETSLSL 659
Qy 684 GKASLARLL-----KLKREQAQKKN 705
Db 660 -QTRFARLLAEYNATQMKVQRLSQLES 686

RESULT 15
I50680
alpha subunit of rod photoreceptor CNG-channel - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: I50680
R:Bonigk, W.; Altenhofen, W.; Muller, F.; Dose, A.; Illing, M.; Molday, R.S.; Kaupp, U.B
Neuron 10, 865-877, 1993
A:Title: Rod and cone photoreceptor cells express distinct genes for cGMP-gated channels
A:Reference number: I50630; MUID:93264082; PMID:7684234
A:Accession: I50680
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-645 <BON>
A:Cross-references: UNIPROT:Q90980; UNIPARC:UPI0000127C21; EMBL:X89599; NID:908852; PID
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotide-
F;431-555/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CAP>

Query Match 18.2%; Score 771; DB 2; Length 645;
Best Local Similarity 34.3%; Pred. No. 3.4e-40;
Matches 173; Conservative 104; Mismatches 196; Indels 32; Gaps 12;

Qy 217 LYLLMLLLVTL--AYNNWCFPLPLVFPYQTDNIHWLIADIICDIYLYDMLFIQPR 274
Db 116 MYNWLFCITPVMYNT--MIARACDFELQNDYLAVWFIVDVSDVIYADM-FVTR 172
Qy 275 LQFVRGGDIIVDSNELKRYRTSTKQLDVASIIIPDICYLFFGFN-PMFRANRMLKYTS 333
Db 173 TGYLEQGLLVKEEQKLAKYKSSIQKLDLFIITDILLYFKLGNYPELRINRLVAR 232
Qy 334 PFEFNHLESIMDKAYIYRVIRTTGYLLFILHINACVYVWASNYEGIGTRRWYDGEKN- 392
Db 233 MFEFFQRTETRTNYPNIFRISNLVMIYIIHWNACVYYSISKAIGFGADTWYYPNTSHP 292
Qy 393 -----EYLCRYVAVRTLITIGLPEPQTLPEIVFOLLNFFSGVFPVSSLIGOMEDVI 445
Db 293 EFARLTRKYVSLYWSLTITTTIGETPPVRDSEYFFVVDLFGVLIFATIVGVGSMI 352
Qy 446 GAATANQNYFRACMDDTIAYNNYSIPIKLQKRVRTWYETWDSQRMDESLLKTLPTT 505
Db 353 SNMNAARAEQAKIDAIKQYMHFENVSKDMKEKVIKWFYLTWNKXADREVLKYLDPK 412
Qy 506 VQLALADNFSIISKVDLFPKCDTQMIYDMLRLKSVLYLPDGFVCKGKEIGKEMYIHK 565
Db 413 LRABIAINVHLEITLTKVRIADFADCEAGLLVELVLKLPQVYSPGDYICRKGDIGREMYIHK 472
Qy 566 HGEVQVLGGPDGKTVLTKAGSVGEISLLAAG--GNERTANVVAHFANLLTLDK 622
Db 473 EGKLAVV-ADGVTVQFVVLSDGSYFGEISILNIGKSGNRRRTANIRISIGYSDLFCLSKD 531
Qy 623 TLQEIIVHYPDSERILMKAR-VLLKQKAKTAEATPPRKDLALLPFPKPEETPKFKTLTG 681
Db 532 DLMEALTEYDPAKAMLEKKGQILMKDGLLDIEVANLGSD-----PKOLEEKV--AYME 583
Qy 682 GTG---KASLARLLKLKREQAQKK 703
Db 584 GMDRLQTKFARL--LAEYDAAQKK 606
```

Search completed: December 21, 2005, 19:58:54

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2005, 19:50:39 ; Search time 48 Seconds  
(without alignments)  
1393.430 Million cell updates/sec

Title: US-09-855-828-1  
Perfect score: 4234  
Sequence: 1 MFKSLTKVKNVKPIGENNEN.....PSAEGGEEVLTIETVEKAKQ 809

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3166.5	74.8	610	2	US-09-949-016-6945
2	2202	52.0	416	2	US-09-949-016-11669
3	1949	46.0	909	2	US-09-538-092-1315
4	857	20.2	698	2	US-09-949-016-10215
5	850	20.1	694	2	US-09-538-092-1351
6	818.5	19.3	690	2	US-09-275-252A-19
7	817	19.3	686	2	US-09-538-092-1037
8	817	19.3	694	2	US-09-949-016-11549
9	655	15.5	575	2	US-09-927-267-1
10	640	15.1	575	2	US-09-927-267-16
11	484	11.4	261	2	US-09-275-252A-18
12	416.5	9.8	256	2	US-09-270-767-46624
13	413	9.8	1203	2	US-09-949-016-6615
14	412.5	9.7	910	2	US-08-997-685A-2
15	412.5	9.7	910	2	US-09-086-436-31
16	412.5	9.7	1109	2	US-09-949-016-10771
17	405.5	9.6	889	2	US-09-949-016-6036
18	399.5	9.4	855	2	US-09-949-016-7263
19	382.5	9.0	749	2	US-08-997-685A-10
20	382.5	9.0	749	2	US-09-086-436-39
21	382	9.0	528	2	US-08-997-685A-4
22	377.5	8.9	504	2	US-09-086-436-33
23	370	8.7	597	2	US-08-997-685A-12
24	368	8.7	506	2	US-08-997-685A-6
25	368	8.7	506	2	US-09-086-436-35
26	368	8.7	597	2	US-09-086-436-41
27	360.5	8.5	960	2	US-09-694-777A-21

28	349	8.2	987	2	US-09-694-777A-22	Sequence 22, Appl
29	347.5	8.2	962	2	US-09-694-777A-24	Sequence 24, Appl
30	345.5	8.2	962	2	US-09-614-480-9	Sequence 9, Appl
31	345.5	8.2	962	2	US-09-694-777A-3	Sequence 3, Appl
32	345.5	8.2	962	2	US-10-422-075-9	Sequence 9, Appl
33	344	8.1	170	2	US-09-358-383C-27	Sequence 27, Appl
34	337	8.0	988	2	US-10-162-012-12	Sequence 12, Appl
35	336	7.9	989	2	US-09-694-777A-23	Sequence 23, Appl
36	335	7.9	988	2	US-09-614-480-2	Sequence 2, Appl
37	335	7.9	988	2	US-10-162-012-5	Sequence 5, Appl
38	335	7.9	988	2	US-10-422-075-2	Sequence 2, Appl
39	334	7.9	319	2	US-09-358-383C-22	Sequence 4, Appl
40	334	7.9	989	2	US-09-694-777A-4	Sequence 4, Appl
41	317.5	7.5	1159	1	US-08-956-242-13	Sequence 13, Appl
42	317.5	7.5	1159	2	US-09-351-215-13	Sequence 13, Appl
43	317.5	7.5	1159	2	US-09-226-012-2	Sequence 2, Appl
44	317.5	7.5	1159	2	US-09-226-012-4	Sequence 4, Appl
45	317.5	7.5	1159	2	US-09-358-383C-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-949-016-6945  
; Sequence 6945, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6945  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6945

Query Match	74.8%	Score	3166.5	DB 2	Length	610			
Best Local Similarity	99.0%	Pred. No.	4e-272						
Matches	609	Conservative	0	Mismatches	1	Indels	5	Gaps	1
Qy	195	MPLTEYLKRIKLPNSIDSYTDRLYLLWLLVLTAYNNWCWFIPRLVFPYOTADNIHWL	254						
Db	1	MPLTEYLKRIKLPNSIDSYTDRLYLLWLLVLTAYNNCCFIPRLVFPYOTADNIHWL	60						
Qy	255	IADIICDIILYDMLFIQPRQFVRGGDIIVDSNELRKHYRTSTKFDQDVASIIPFDICY	314						
Db	61	IADIICDIILYDMLFIQPRQFVRGGDIIVDSNELRKHYRTSTKFDQDVASIIPFDICY	120						
Qy	315	LFFGFNPFANRMLKYTSFFFNHLESIMDKAYIVIRTTGYLLFILHINACVYWA	374						
Db	121	LFFGFNPFANRMLKYTSFFFNHLESIMDKAYIVIRTTGYLLFILHINACVYWA	180						
Qy	375	SNYEGIGTRWYDGEGNEYLRCCYVAVRTLITIGGLEPOTLFEIVFOLLNFFSGVVF	434						
Db	181	SNYEGIGTRWYDGEGNEYLRCCYVAVRTLITIGGLEPOTLFEIVFOLLNFFSGVVF	240						
Qy	435	SSLIGQMRDVTGAATANQNYFRACMDDTIAYMNNYSIPKLQKRVTRWYETWDSQRLMD	494						
Db	241	SSLIGQMRDVTGAATANQNYFRACMDDTIAYMNNYSIPKLQKRVTRWYETWDSQRLMD	300						
Qy	495	ESDLLKLTPTTVQLALADVNFISIISKVDLFGKCDTQMIYDMLLRKLSVLYLPGDFVCKK	554						

Db 301 ESDLKLTPTTQVLAALDVNFSSILSKVDLFGKCDTQMIYDMLRLKSLVLYLFGDFVCKK 360  
QY 555 GEIGKEMYIIKHGEVQVLGGPDGKTVLTKAGSVFGEISLLAAGGNGRRTANVAHGFA 614  
Db 361 GEIGKEMYIIKHGEVQVLGGPDGKTVLTKAGSV-----LLAAGGNGRRTANVAHGFA 415  
QY 615 NLLTLDKKTQLBILVHYPDSERILMKKARVLLKQAKTAEATPPRKDLALLFPKKEETPK 674  
Db 416 NLLTLDKKTQLBILVHYPDSERILMKKARVLLKQAKTAEATPPRKDLALLFPKKEETPK 475  
QY 675 LFKTLLGGTGKASLARLLKKEQAAKXENEGEGEKEKEDKQKEDKQKEDKXG 734  
Db 476 LFKTLLGGTGKASLARLLKKEQAAKXENEGEGEKEKEDKQKEDKQKEDKXG 535  
QY 735 KENEDKXGKEPEEKPLDRPECTASPIAVEEPHPSVRRTVLPRTGTSRQSLIISMAPSARG 794  
Db 536 KENEDKXGKEPEEKPLDRPECTASPIAVEEPHPSVRRTVLPRTGTSRQSLIISMAPSARG 595  
QY 795 GBEVLTIKKEKAKQ 809  
Db 596 GBEVLTIKKEKAKQ 610

## RESULT 2

US-09-949-016-11669  
; Sequence 11669, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11669  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11669

Query Match 52.0%; Score 2202; DB 2; Length 416;  
Best Local Similarity 99.8%; Pred. No. 9.2e-187;  
Matches 412; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 114 POKKPPAAVINEYADAQLHNLVKGMORTALYKXKLVGGDLSSEASPTAKPTAVPPV 173  
Db 1 POKKPPAAVINEYADAQLHNLVKGMORTALYKXKLVGGDLSSEASPTAKPTAVPPV 60  
QY 174 KESDDKPTHEYRLWLFVKKMPLEYLKRILKLPNSIDSYTDRLLVLLVLLVTLAYNWC 233  
Db 61 KESDDKPTHEYRLWLFVKKMPLEYLKRILKLPNSIDSYTDRLLVLLVLLVTLAYNWC 120  
QY 234 WFIPLRLVPPYQTDNIHNLWLIADICDIILYDMLFIQPRLOFVRGGDIIVDSNELRKH 293  
Db 121 WFIPLRLVPPYQTDNIHNLWLIADICDIILYDMLFIQPRLOFVRGGDIIVDSNELRKH 180  
QY 294 YRTSTKFOLDVASIIPEDICYLFFGPNPMFRANMLKYSFEEFNHLESIMDKAYIYRV 353  
Db 181 YRTSFQFDVASIIPFDICYLFFGPNPMFRANMLKYSFEEFNHLESIMDKAYIYRV 240  
QY 354 INTTGYLLFILHINACVYVWASNYEGIGTRVWYDGEKNEYLRCYVWAVRTLITIGGLPE 413  
Db 241 INTTGYLLFILHINACVYVWASNYEGIGTRVWYDGEKNEYLRCYVWAVRTLITIGGLPE 300

QY 414 PQTLEIIVFQLNLPFSGVFVSSLLIGQMRDVIIGAATANQNYPFRACMDDTIAYMNNYSIPK 473  
Db 301 PQTLEIIVFQLNLPFSGVFVSSLLIGQMRDVIIGAATANQNYPFRACMDDTIAYMNNYSIPK 360  
QY 474 LVQKEVRTWYETWDSQRMDESLLKLTPTTVQALALADIVNFSIISKVDLFPK 526  
Db 361 LVQKEVRTWYETWDSQRMDESLLKLTPTTVQALALADIVNFSIISKVDLFPK 413

## RESULT 3

US-09-538-092-1315  
; Sequence 1315, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuraPatSeqformatter Version 0.9  
; SEQ ID NO 1315  
; LENGTH: 909  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number Q14028  
US-09-538-092-1315

Query Match 46.0%; Score 1949; DB 2; Length 909;  
Best Local Similarity 46.2%; Pred. No. 9e-164;  
Matches 403; Conservative 122; Mismatches 238; Indels 110; Gaps 18;

QY 14 IGENNENQSSRRNEEGSHPSNQSOQTAAOE--ENKGE---EKSUKTKSTPVTSEPHTN 68  
Db 67 VGEAEKAEKAEKAE 126  
QY 69 IQDK-----LSKKNSSGDLTTNPDPONAAEPTGTVPQEKEMDMPGKEGN----- 112  
Db 127 VEDTDADSCPLMAEENPPS--TVLPSPSPAKSDTLIVPSSASGTHRKLPSEDDAEELK 184  
QY 113 --SPONKP-----PAAP-----VINEYADAQLHNLVKRMORTALY 146  
Db 185 ALSPAESPVAWSDPPTTPKDTGQDRAASTASTNSAIIND---RLQELVKLFKERTKV 240  
QY 147 KKKLVEGDL---SPEASPTAKPTAVPPVKESDDKPT--EHYRLLWFKVKKMPLTEY 200  
Db 241 KEKLIDPDTVDSSESPKPSAKKAPAPADTKPAEAEPEVEEHCMLCCFKPRPWKKY 300  
QY 201 LKRILKLPNSIDSYTDRLLVLLVTLAYNWCNCFIPRLRVFPYQTDNIHNLWLIADIC 260  
Db 301 ---QFQSIDPLTNLMVLMVLMVLMVLMVLMVLMVLMVLMVLMVLMVLMVLMVLMVLMV 356  
QY 261 DIILYDMLFIQPRLOFVRGGDIIVDSNELRKHRTSTKFOLDVASIIPFDICYLFFGPN 320  
Db 357 DLIIYDMLFIQPRLOFVRGGDIIITDKDMRNLYKSRFRKMDLLSLPLDPLFLYLVKGVN 416  
QY 321 PMFRANMLKYSFEEFNHLESIMDKAYIYRVIRTTTCYLLFILHINACVYVWASNYEGI 380  
Db 417 PLLRLPCKKFMFEEFNHLESILSKAYIVRVIRTTTCYLLFILHINACVYVWASNYEGI 476  
QY 381 GTTRWYDGEKNEYLRCYVWAVRTLITIGGLPEPQTLFEIVFQLLNFPFSGVVFSSLLIG 440  
Db 477 GSTHWYDGVNSYIRCYFAVKTLITIGGLPDKTLFEIVFQLLNFPFSGVVFSSLLIG 536  
QY 441 MRDVIGAATANQNYPFRACMDDTIAYMNNYSIPKLVQKVRTWYETWDSQRMDESLLK 500

Db 537 MRDVGAATAGQTYRSCMDSTVKYKMFYKPKSVQNRVKTWYETWHSQGLDSELMV 596  
QY 501 TLPTTQLALADVNFSIIISKVDLFGCDTQIMYDMLRLKSVLYLPDGFVCKKGIGKE 560  
Db 597 QLPDKWRLDALDVNINYSKVALFGCDQIMFDMLKRLSVVLYLPNDYVCKKGIGRE 656  
QY 561 MYIIKHGVEQVGLGGPDGTVLTKAGSVFGISLLAAGGNNRTANVVAHGFANLLTLD 620  
Db 657 MYIIQAGQVQLGGPDGKSVLTKAGSVFGISLLAAGGNNRTANVVAHGFANLLTLD 716  
QY 621 KKTQLBILVHPDSEILMKARVLKOKAKTAEATPPKOLALLPFPKEETPKLFTLL 680  
Db 717 KCOLNEILVHPESQKLLRKARRMLRSNNK-----PKBEKSVLLILPPRAGTPKLFNAAL 771  
QY 681 GGTGKAS-----LARKLLKREAAQKENSEGEGEGEKENEKEDKOK 728  
Db 772 AMTGKMGKAGKGLAHLARLKLALALEAAKHEEL---VEQAKSSQDVKEGESAAP 828  
QY 729 ENEDKQKE-NEDKDKGREPEEKPLDRPECTASPIAV-----EPEHVSVRTVLP 776  
Db 829 DQHTPEKATDPPAPRTPEPP-GSPSPSPASLGSCGEGEGEAGPEEHSVR-----882  
QY 777 RGTSRSLIISMAPSAGGSEVLITIEVKEKAKQ 809  
Db 883 -----ICMSPGPEGQILSVKMPERE 906

## RESULT 4

US-09-949-016-10215  
; Sequence 10215, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10215  
; TYPE: PR1  
; ORGANISM: Human  
US-09-949-016-10215

Query Match 20.2%; Score 857; DB 2; Length 698;  
Best Local Similarity 29.9%; Pred. No. 5.2e-67;  
Matches 226; Conservative 153; Mismatches 272; Indels 106; Gaps 22;

QY 19 ENEQSRNRNGBSPHNSQSQTTAQEENKGEKSLTKTPTVTSEEPHTNIQDKLSKNS 78  
Db 1 QTEKXAKINTQYSHPS-----RTHLKVK-----TSRDLNRAENGLSRAHS 41  
QY 79 SGDLTTNPDQNAEPTGTVPQKEMDPGKGNPSQNKPPAAPVINEYADAQLNLVXR 138  
Db 42 SSBETS-----SVLPQ-GIAMEIRGLADSGGSGFTGQ-----GIARSLRILFL 83  
QY 139 MRQRTALYKXKLVGDLSSP-----EASPTAKTAPVPPKESDDKTEHYHLLW 189  
Db 84 LRRWAARHVHQDQGPSPFDRFRGABLKVEYSSQESNAQANVGQBPADRG-----BSAW 138  
QY 190 FKVKMPL-----TEYLKEIKLPNSI-DSYTDRLYLLLVLTAYNNWCWPIPL 238  
Db 139 -----PLAKCNTWTSNNTEBEKTKKDALVDPDPSNLYRWLTATLALPVFNWYLLIC 192

QY 239 RLVPYQYADNTHVYLIADIICDIILYDMLFIQPRLOFVRGDIIVDSNELRKHYRTST 298  
Db 193 RACFDELOSEYLMWLVLVDYSADVLYVLDVL-VRARTGFLBOGLMVSNTNRNLWOHYKTTT 251  
QY 299 KFQLDVASIIPEDICVLPFGFN-PMFRANRMLKYTSFFEFNHHLESIMDKAYIYRVIRTT 357  
Db 252 QPKLDVLSVPTDLAYLVKGTNYPEVRNRLKLSRLEFFDPTETRTTNYPMFRIGNLV 311  
QY 358 GYLLFILHINACYWASNYEGIGTRVYDG-----EGNEYLRCYVWAVRTLTITIG 409  
Db 312 LVILIIHWNACIYFAISKFIGFGTDSWVYPNISIPHEHRLSRKYIYSLYWSLTFLTITIG 371  
QY 410 GLPEQTLFEIVFQLLNFPESGVFSSLLIGORDVIGATANQNYFRACMDDTIAYMNY 469  
Db 372 ETTPPPVKDEYLVVVVDVFLVGLVIFATIVGNVGSIMSNNASRAEFOAKIDSIKQYMQFR 431  
QY 470 SIPKLQKRVRTWYEYTWDSORMLDESLLKTLPTTVOLALAI DVNFSIISKVDLFGCD 529  
Db 432 KYTKOLETRVIRWFDYLMANKKTVDKEVLKSLPDKLKAELAINVHLDTLKKVRI FODCE 491  
QY 530 TQMIYDMLRLKSVLYLPDGFVCKKGIGKEKEMVYIIKHGEVOVLGGPDGTVLTKAGSV 589  
Db 492 AGLLVELVLKLRPTVPSPGDIYCKGDIKEMVYIINEGKLAVV-ADGGVTOFVVLSDGSY 550  
QY 590 FGEISLLAAGG---GNRTANVVAHGFANLLTLDKKTLOEILVHYPDSEILMKKAR-VL 645  
Db 551 FGEISILNITKSGSGNNRTANIRSIGSDLFCLSKDDLMEALTEYPEAKKALEEGROIL 610  
QY 646 LKQK---AKTAEATPPKDLALLFPKKEETPKLFTLLGCTGKASLARLL-----KL 694  
Db 611 MKDNLIDEELARAGADPKDL-----EKVEQLGSSL--DTLOTRFARULLAEYNATOMKM 662  
QY 695 KREQAQKXENEGEGEE---EGKENEDKOKENEDKOK 728  
Db 663 KQRLSQLESQVKGKDKPLADGVEVPDATK-TEDKQO 698

## RESULT 5

US-09-538-092-1351  
; Sequence 1351, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CurapSeqFormatter Version 0.9  
; SEQ ID NO 1351  
; LENGTH: 694  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number Q16281  
US-09-538-092-1351

Query Match 20.1%; Score 850; DB 2; Length 694;

Best Local Similarity 30.0%; Pred. No. 2.2e-66;  
Matches 225; Conservative 149; Mismatches 269; Indels 106; Gaps 22;

QY 27 NEEGSHPSNOSQQTAAQENKGEKSLTKTSTPVTSEEPHTNIQDKLSKNSGDLTTNP 86  
Db 5 NTQVSHPS-----RTHLKVK-----TSRDLNRAENGLSRAHSSEETS-- 43  
QY 87 DPQNAEPTGTVPQKEMDPGKGNPSQNKPPAAPVINEYADAQLNLVXMRQRTALY 146

Db 44 ---SVLQP-GIAMETRGADSGQGSFTG-----GIARLSRLIFLLRWAARH 87  
Qy 147 KKKLVGDLSSP-----EASPTAKPTAVPPVKESDDKPTHEHYVLLAFKVKKMP 197  
Db 88 VHQDQGPSFDFRFGAELKEVSSQESNAQNVGQEPADRG-----RSAM-----PL 136  
Qy 198 -----TEYLKRIKLPNSI-DSYTDRLYLMLLVLTLAYNMNCFPIRLVFPYQT 246  
Db 137 AKCNTWTSNTEBEKTKKDAIVDPSSNLVYRLTAIPALPVYNNWLLICRACFDELQ 196  
Qy 247 ANIHVWLADIICDIILYVDMFLTPQRLQFVRGGDIIIVDSNELKHYTSTKFKQLDVAS 306  
Db 197 SEYMLVLDYSADVLVLDVL-VRAQTGLEQGLMVSDTNRLMWHYKTTTQFKLDVLS 255  
Qy 307 IIPFDICYLFFGFN-PMFRANRMLKYTSFPEFNHLESIMDKAYIVRVIRTTGYLLFILH 365  
Db 256 LVPTDLAYLVKGTNYEVPFRNLKFSRLEFPDRTETRTNYPNMRIGNLVLIILIH 315  
Qy 366 INACVYVNASYEGIGTTRVYDG-----EGNEYLRCYVAVRVLITIGGLPEPOTL 417  
Db 316 MNACIVFAISKFIGTDSWVYPNISIPEHGRLSRKYIYSLYWSLTLTLTIGTTPPVKD 375  
Qy 418 FEIVFOLLNPFSGVVFSSLIQMRDVIGAATANQNYFRACMDDTTAYMNNVSIPLKQVK 477  
Db 376 EYFLFVVVDPLVGLVLPATVIGNVGSMISNMNASRAEFOAKIDSIRKQYQFRKVTKDLET 435  
Qy 478 RVRTVEYTWDSQRMLEDLLKTLPTTVQLALAIQVNFSTISKVDLPFKGCDTQMIYDML 537  
Db 436 RVIRFDYLWANKTVDEKEVLSLDPDKAEIAINVHLDTLKKVRFQDCEAGLLVELV 495  
Qy 538 LRLKSVLYLPDGFVCKKKGIGKEMYIIKHGEVQVLGGPGDTKVLVTLKAGSVFGEISLLA 597  
Db 496 LKLRPTVSPGDYICKKGDIKEMYIINEGKLAVV-ADGVTQFVLSGDSYFGEISILN 554  
Qy 598 AGG---GNRRTANVAHGFANLLTLDKKTLOEILVHYPSERILMKAR-VLLKQK---A 650  
Db 555 IKGSGSGNRRTANIRSIGSDLFCLSKDDLMALTEYPAKKALEBKGRQILMKDLIDE 614  
Qy 651 KTAETPPPKDLALLPPPEETPKLPKTLGGTGKASLARLL-----KLKREQAQK 702  
Db 615 ELARAGADPKDL-----BEKVEQLGSSL--DTLQTRFARLLAEYNATQMKQRLSQLE 666  
Qy 703 KENSEGEE---EGKENEDKQENEDKQK 728  
Db 667 SQVKGCGDKPLADGEVPGDATK-TEDKQK 694

RESULT 6

US-09-275-252A-19  
; Sequence 19, Application US/09275252A  
; Patent No. 6641997  
; GENERAL INFORMATION:  
; APPLICANT: MacKinnon, Roderick  
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With  
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation  
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof  
; FILE REFERENCE: 018512-002901US  
; CURRENT APPLICATION NUMBER: US/09/275,252A  
; CURRENT FILING DATE: 1999-03-24  
; PRIOR APPLICATION NUMBER: US 09/045,529  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: US 09/054,347  
; PRIOR FILING DATE: 1998-04-02  
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307  
; PRIOR FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-275-252A-19

Query Match 19.3%; Score 818.5; DB 2; Length 690;  
Best Local Similarity 30.0%; Pred. No. 1.3e-63;  
Matches 221; Conservative 128; Mismatches 262; Indels 125; Gaps 22;  
Qy 5 LTKVNVKPIGNNENQSSRRNEEGSHPS-----NOSQOTTAQENKKEE-----KSL 53  
Db 17 VTMNVIVP-----DIEKEIRRMENGACSSPSEDDSDSAVTSSESENPHARGSFYKSL 71  
Qy 54 KTKSTPVTSEBPHTHIQDKLSKNSGDLTTPNDPQNAAEPTGTVPQEKEMDPGKEGPN 113  
Db 72 R-KGGSQREQLPGAIAIFNVNNS-----NKD-----QEPEEKKKKKKKSKDDKNE 121  
Qy 114 POKKPPAAPVINEYADAQLHNLKRMORTALYKKKLVGEGDLSSEASPTAKTAVPPV 173  
Db 122 NKNDPE-----KKKKKKKKKK-----E 141  
Qy 174 KESDDKPTHEYVRLWFKVKMPLTEYLKRIKLPSIDSYTDRLYLMLLVL AYNW 231  
Db 142 EKSQKKEHH-----KKEVV-----IDP-SCNTYNNMLFCITLPMYNNW 140  
Qy 232 NCWFIPLRVFPYQADNIHYWLIADIICDIILYVDMFLTPQRLQFVRGGDIIIVSNEUK 291  
Db 181 T--MVIARACFDELOSDYLEYWLILDYVSDIVYLIDM FVRTRTGYLEQGLLVKEELKI 237  
Qy 292 KHYRTSTKFPOLDVASIIPFDICYLFFGFN-PMFRANRMLKYTSFPEFNHLESIMDKAYI 350  
Db 238 NKYKSNLOPKLDVLSLIPDTDLFLKLGWNYPEIRLNRLLRFSRMPFEFFORTETRTNPN 297  
Qy 351 YRVIRTTGYLLFILHINACVYVNASYEGIGTTRVY-----DGE---GNEYLRCYVAV 402  
Db 298 PRISNLVMYVIIHWNACVFYSISKALCFGNDTWVYPDINDPEFRLARKVYVSLYWS 454  
Qy 403 RFLITIGGLPEPOTLFEIVFOLLNPFSGVVFSSLIQMRDVIGAATANQNYFRACMDDT 462  
Db 358 LTLTTIGTTPPVRDSEYFVVVDLGLVLPATVIGNIGSMISNMNARAEFOARIDAI 417  
Qy 463 IAYMNNYSIPKLVQKRVTRTWEYTWDSQRMLEDLLKTLPTTVQLALAIQVNFSTISKV 522  
Db 418 KQYMFPRVSKDMKRVIKWFDYLTNNKTVDEKEVLPDKLRAEIAINVHLOTLLKV 477  
Qy 523 DLFKGCDTQMIYDMLRLKSVLYLPDGFVCKKKGIGKEMYIIKHGEVQVLGGPGDTKVLV 582  
Db 478 RIFADCEAGLLVELVLQPVYSPGDYICKKGDIKGRMYIIKEGKLAVV-ADGVTQFV 536  
Qy 583 TLKAGSVFGEISLLAAGG---GNRRTANVAHGFANLLTLDKKTLOEILVHYPSERILM 619  
Db 537 VLSGSGTTFGEISILNIGKSGAGNRRTANIKSIGSDLFCLSKDDLMALTEYPAKTMLE 596  
Qy 640 KVARVLLKQK----KTAETPPPKDLALLPPPEETPKLPKTLGGTGKASLARLL--- 692  
Db 597 EKGQILMKDGLLDLNIANAGSDPKDLEEKVTRMEGSDVLLQT-----RPARILAEY 648  
Qy 693 -----KLKREQAQK 703  
Db 649 ESMQKQLKQLTKVEK 664

RESULT 7

US-09-538-092-1037  
; Sequence 1037, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965

```

; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratPatSeqformatter Version 0.9
; SEQ ID NO 1037
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P29973
US-09-538-092-1037

Query Match
Best Local Similarity 19.3%; Score 817; DB 2; Length 686;
Matches 226; Conservative 129; Mismatches 266; Indels 136; Gaps 24;

QY 5 LTKVNVKPIGENNENQSSRRNEEGSHPS-----NOSQOTTAQENKGE-----KSL 53
DB 13 VTMPNVIVP-----DIEKEIRRMENGACSSFSDDDSASTSESENENPHARGSFYSKSL 67
QY 54 KTKSTPTVTEEPHTNIQDKLSKNSGDLTTPDPQNAAEPTGTVPEQKEMDPGKGPNS 113
DB 68 R-KGGSQREQLPGAIALFNVNSS-----NKD-----QEPBEKKKKKKSKSDKNE 117
QY 114 POKKPPAAPVINEYADAQLHNLVKRMORTALYKKLVGDLSSPEASPTAKPTAVPV 173
DB 118 NKNDPE-----KXKKKKKKKKKK-----E 137
QY 174 KESDDKPTHEYRLLMFKYKMPLETEYLKRIKLPNSIDSYDRLYLMLLLVLT--AYNW 231
DB 138 EKSDDKKEE-----KKEVVV-----IDP-SGNTYYNWLFCITLPMYNNW 176
QY 232 NCWFILRLVFPYOTADNIHYMLIADIICDIIYLYDMLFIQPLQFVGGDIIVDSNELR 291
DB 177 T--MVIARACFDELQSDYLEYMLIDYVSDIVYLIDM-FVTRTGYLEOGLLVKEELKLI 233
QY 292 KHYRTSTKFOLDVASIIPFDICVLPFGFN--PMFRANRMLKYTSFFEFNHHLESIMDKAYI 350
DB 234 NKYSNLQPKLDVLSLIPTDLYFLGWNYPEIRLNLRLFRSMFEFFORTETRTNYPNI 293
QY 351 YRVIRTTGYLLFILHINACVYVWASNYEGITTRWY-----DGE-----GNEYLCRYWAV 402
DB 294 FRISNLVMYIVIIHWNACVYSISKAIGFGNDTWVYVINDPDPFGLRLARKYVYSLWST 353
QY 403 RTLITIGLPEPOTLPEIVFOLLNFGVVFSSLGOMRDVIGAATANONYFRACMDOT 462
DB 354 LTLTTIGETPPPPRDSEYVFWVDFLIGVLIPATIVGNISMSNNNAARAEQARIDAI 413
QY 463 IAYMNNYSIPKLQKRVRTWYETWDSQRMLESDDLKTLPTTVQLALADIVNFSISKV 522
DB 414 KQYMHFRNVSKDMEKRVIKWFDYLTNKTVDKEVLYKLPDKLRAEIAINVHLDLTKKV 473
QY 523 DLFKGCOTQMIYDMLRLKSVLYLPGDFVCKKGEIGKEMYIIKHGEVQVVGPGDGTKVLV 582
DB 474 RIFADCEAGLLVELVLKLPQVYSPGDYICKKGDI GREMYIIKEGLAVV-ADGVTQFV 532
QY 583 TLKAGSVFGISILAAAG---GNRRANVVAHGFANLLTLDKKTLQBIHVYDSEIRLM 639
DB 533 VLSGYSVFGISILNIKSGKAGNRRTANIKSIGYDLFCLSKDLMALTEYDPAKTMLE 592
QY 640 KQARVLLKQKA-----KTAEATPPRKLALLFPKKEETPKLTKLGGTGKASLARLL--- 692
DB 593 EKGQILMKDGLLDLNIANAGSDPKLEEKVTRMEGSDVLLQT-----RFAIILAEY 644
QY 693 -----KLKREQAQK-----ENSEG-GEEG 713
DB 645 ESMOQKLQRLTKVEKFLKPLIDTEFSSIEGPGAEG 681

RESULT 8
US-09-949-016-11549
; Sequence 11549, Application US/09949016

```

```
QY 693 -----KLKEQAQK-----ENSEG-GEERG 713
Db 653 ESMQKQLKQLTKVFKFLPLIDTFESSIEGPGAESG 689

RESULT 9
US-09-927-267-1
; Sequence 1, Application US/09927267
; Patent No. 6933147
; GENERAL INFORMATION:
; APPLICANT: Crech, Christopher D.
; APPLICANT: Jegla, Timothy J.
; APPLICANT: ICAGEN, Inc.
; TITLE OF INVENTION: Channel
; FILE REFERENCE: 018512-006510US
; CURRENT APPLICATION NUMBER: US/09/927,267
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/226,253
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: cyclic nucleotide-gated cation channel 2B (CNG2B)
US-09-927-267-1

Query Match 15.5%; Score 655; DB 2; Length 575;
Best Local Similarity 29.2%; Pred. No. 3.3e-49;
Matches 165; Conservative 121; Mismatches 241; Indels 38; Gaps 14;

QY 202 KRIKPNISDSTDRILYLLWLLVTLAYNNWCFPLRLVFPYQTADNIHYWLIADIICD 261
Db 19 KARKLPVLDPGSD-YYWWLNTWVFPVYNNLIILVCRAFPDQLQHGVLVWLVLDYDTS 77

QY 262 IIVLYDMLFIQRLQFVRGDIIVDSNELRKHVRTSTKFDQDVASIIIPDICYLFFGFN 320
Db 78 LIYLLDMV-VRHTGFELEGILVVDKGRISRYVRTWSFFLDASLMPDVTYVRLGPH 136

QY 321 PMFRANRMLKYTSFFEFNHLESIMDKAYIYRVIRTGGYLLFILHINACVYVWASNYEGI 380
Db 137 PTLRLNRFRLAPLPEAFDTRTAYPNAFRIAKMLYIFVVIHNSCLYFALSRYLGF 196

QY 381 GTRWYDDEG-----NEVLCRYVAVRTLITIGGLPEPQTLFEIVFOLLNFFSGV 432
Db 197 GRDAWYPPDPAQGFRLRQYLSFYFSTLITITVGDTPPPAREBEYLFMVGDFLAYM 256

QY 433 VFSSILIGMRDVI-GAATANONYFRACMDDTIA--YMNYSIPKLQKRVRTWYETWDS 489
Db 257 GFATINGSMSVYNNMTAAAFY----PDHALVKYKMLQHVNRKLERVIDWYQHLQIN 313

QY 490 QRMLDESLLKTLPTTVQLALADIVNFSIISKVDLFGKCDTQMIYDMLRLKSLVLYLPGD 549
Db 314 KQWTEVAILOLQHLPERLRAEVAVSVHLSTLSRVQIFQNCESALLEELVLKLPQTSVCE 373

QY 550 FYCKGEIGEKEMIIKHGEVOVLGGPDGKVLVTLKAGSVFGEISLLAAGG---GNRRTA 606
Db 374 YVCRIGDIGQEMIIIREGOLAVV-ADGGITQYAVLGAGLYFGEISINIKGNMSGNRRTA 432

QY 607 NVVAFGANLLTLDDKTLQELIVHYPDSERILMKKAR-VLLK-----QAKTA--EA 655
Db 433 NTKSGYSLDFCLSKEDLREVSEYQATIMEKGREILLKMNKLDVNAEAEIALQEA 492

QY 656 TPPRKDLALLPPPKKEPTPLFKTLTGGTGKASLARLLKLRQEAQAKK----ENSEGSEE 711
Db 493 TESR--LRGLDQQLDQLTKFARLLAELESSALKIAYRIERLEWQTPWMPEDLAEADD 550

QY 712 EKENEDKOKENEDKOKENEDKGE 736

QY 693 -----KLKEQAQK-----ENSEG-GEERG 713
Db 653 ESMQKQLKQLTKVFKFLPLIDTFESSIEGPGAESG 689

RESULT 10
US-09-927-267-16
; Sequence 16, Application US/09927267
; Patent No. 6933147
; GENERAL INFORMATION:
; APPLICANT: Crech, Christopher D.
; APPLICANT: Jegla, Timothy J.
; APPLICANT: ICAGEN, Inc.
; TITLE OF INVENTION: Channel
; FILE REFERENCE: 018512-006510US
; CURRENT APPLICATION NUMBER: US/09/927,267
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/226,253
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: rat cyclic nucleotide gated cation channel OKNC2
US-09-927-267-16

Query Match 15.1%; Score 640; DB 2; Length 575;
Best Local Similarity 29.1%; Pred. No. 7.1e-48;
Matches 157; Conservative 114; Mismatches 224; Indels 44; Gaps 14;

QY 218 YLLWLLVTLAYNNWCFPLRLVFPYQTADNIHYWLIADIICDIIYLYDMLFIQRLQF 277
Db 34 YWWLNTWVFPVYNNLIIVVCRAFPDQLQHGVLVWLVLDYDTSLLYLLD-IGVRFHTCF 97

QY 278 VRGSDIIVDSNELRKHVRTSTKFDQDVASIIIPDICYLFFGFN-PMFRANRMLKYTSFFE 336
Db 93 LEQGLVVDKGMIAISRYVRTWSFFLDASLVPTDAAYVOLGPHIPTLRNRFVRPLRFE 157

QY 337 FNHLESIMDKAYIYRVIRTGGYLLFILHINACVYVWASNYEGICTTWWYDGE- 491
Db 153 AFDRTETRTAYPNAFRIAKMLYIFVVIHNSCLYFALSRYLFGORDAWVYDPAQGF 212

QY 392 ---NEVLCRYVAVRTLITIGGLPEPQTLFEIVFOLLNFFSGVVFSSILIGMRDVI 447
Db 213 RLRRQYLSFYFSTLITITVGDTPLPDREBEYLFMVGDFLAVMGFATIMSSSVIYNN 272

QY 448 ATANONYFRACMDDTIA--YMNYSIPKLQKRVRTWYETWDSQRMLEDSDLLKTLPTT 505
Db 273 NTADAAFY---PDHALVKYKMLQHVNRKLERVIDWYQHLQINKKMTNEVALQHLPER 129

QY 506 VOLALADIVNFSIISKVDLFGKCDTQMIYDMLRLKSLVLYLPGDFVCKKGEIGKEMYIIK 565
Db 330 LRAEVAVSVHLSTLSRVQIFQNCESALLEELVLKLPQTSVGEYVCRKDGIGREMYIIR 389

QY 566 HGEVOVLGGPDGKVLVTLKAGSVFGEISLLAAGG---GNRRTANVVAHGFANLLTLDDK 622
Db 390 EQGLAVV-ADGGITQYAVLGAGLYFGEISINIKGNMSGNRRTANIKSLGYSDLFCLSK 448

QY 623 TLQELIVHYPDSERILMKKAR-VLLK-----QAKTA--EATPPRKDLALLPPPKKE 671
Db 449 DLREVLSVEYVQAQVMEEGREILLKMNKLDVNAEAEIALQEAESR LKGLDQQLD 506

QY 672 TPKLFKTLTGGTGKASLARLLKLR -----EQAQKKNSESGEEGKENEK 714
Db 507 LQTKFARLLAELESSALKIAYRIERLEWQTPWMPEDMEADDEAEPEGTSGDEGK 565

RESULT 11
US-09-275-252A-18
; Sequence 18, Application US/09275252A
; Patent No. 6641997
```

```

; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact with
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/09/275,252A
; CURRENT FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-275-252A-18

Query Match      11.4%; Score 484; DB 2; Length 261;
Best Local Similarity 37.2%; Pred. No. 1.5e-34;
Matches 97; Conservative 58; Mismatches 94; Indels 12; Gaps 4;

QY 370 VYWAASNYEGIGTRWVY----DGE----GNEYLRCYVAVRTLIITIGGLPQPOTLFEIV 421
DB 1 IYVAISKSGFGVDWTWYFNITDPEGYLAREVIYCLYVSTLTITIGTTPPVKDEEVL 60
QY 422 FQLLNFFSGVVFSSLIGQMRDVIGAATANQNVFRACMDDTIAYMNNYSIPKLQKRVRT 481
DB 61 FVIFDELIGVLIFATVGNVGMISNNATRAEFOAKIDAVKHYMQFRKVSKGMEAKVIR 120
QY 482 WBYETWDSQRMDESLLKTLPTVQLALAIQVNFISIKVDLFKCDTQMIYDMLLRK 541
DB 121 WFDYLTNKTVDREIRILKNLPAKRAEINAVHLSLTKKKVIRIFHDCBAGLVLVLKLR 180
QY 542 SVLYLPQDFVCKGEGEKEMWIIKHGEVQVGLGGPDGTVLTLKAGSVFGEISLLAAGG- 600
DB 181 PQVFSFGDYICRGDIOGEMWIIKHGEVQVGLGGPDGTVLTLKAGSVFGEISLLAAGG- 600
QY 601 --GNRRATANVAHFANLITL 619
DB 240 KMGNNRRATANIRSIGSDLFCL 260

RESULT 12
US-09-270-767-46624
; Sequence 46624, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46624
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-46624

Query Match      9.8%; Score 416.5; DB 2; Length 256;
Best Local Similarity 36.8%; Pred. No. 1.4e-28;
Matches 78; Conservative 53; Mismatches 80; Indels 1; Gaps 1;

QY 253 WLADIICDIYLYDMLFIQRLQFVRGGDIIVDSNELRKHYRTSTKTFOLDVASIIPFDI 312

```

```

DB 45 WLACDFCADIYLLDQVFFKHRVMYLFEGFWKKNLTKNMRKQLQKLDLALLPLEL 104
QY 313 CYLFFGFMFMF-RANRMLKYTSFFEFNHHLESIMDKAYIVRVIRTTGYLLFILHINACVY 371
DB 105 LYFKLGTQAVMLRPPFPFKIQSFWEVPRLLDRVISSPHFVRVAKLTLYMLYMIHITAALY 164
QY 372 YWASNYEGITRWVYDGEHNEYLRCYVAVRTLIITIGGLPQPOTLFEIVFOLLNFFSGV 431
DB 165 YAYSQYQGLGNRWVFSKGHPYVRCFAFATKTATSIGKNPKPERQGEYVFMTVAWLMGV 224
QY 432 FVFSSLIGQMRDVIGAATANQNVFRACMDDTI 463
DB 225 FVALLIGQIRDIISTATRNKHYRQLEDEL 256

RESULT 13
US-09-949-016-6615
; Sequence 6615, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6615
; LENGTH: 1203
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6615

Query Match      9.8%; Score 413; DB 2; Length 1203;
Best Local Similarity 23.9%; Pred. No. 3.2e-27;
Matches 152; Conservative 109; Mismatches 273; Indels 101; Gaps 25;

QY 62 SBEPTNIQDKLSKN--SSGDLTTNPD---PQNAEPTGTVPQEKEMDPGKE-----GP 111
DB 115 SGSSGHHLHDSAEERRLIAEGDASPCEDRTPPCLAAEP-----RPGASAPAAASP 165
QY 112 NSPQKPPAAPVINEYADALQHLNVLKRMQRTAL-YKKKLVEGDLSSPASPOKAPTAV 170
DB 166 PPPQPPQPASASCEQPSVD-----TAIKVEGGAAGADOILPEAEVRLGAGFM 214
QY 171 PPVKESDDKPTHEYHYRLLWFKVKMKPLTEYLKRIKLPN--SIDSYTD-RLYLMLLVL 227
DB 215 QRQFGAMLQGVGNKFSRLRMFGSOKAVEREQ-ERVKSGAGFWIHPYSDFRPY--WDLTMLL 271
QY 228 AVNWCNWFPLRLVFPYQYADNIHYWLIADIICDIYLYDMLFIQRLQFVRG-----G 281
DB 272 LMVGNLIIIPVGITF--FKDENTTPWVFNVSVDTFLLDLV-----LNFRTGIVVEDNT 324
QY 282 DIIVDSNELRKHYRTSTKTFOLDVASIIPFDICVL-----FFGNPMF 323
DB 325 EILDPQRIMKYLKSW-FMVDPFISSIPVDYIFLIVETRIDSEVYKTRALRIVRTKIL 383
QY 324 RANRMLKYTSFFEFNHHLESIMDKAY-----IYRVIRTTGYLLFILHINACVYASNYE 378
DB 384 SLURLRLRLIRYIHOWBEEIFHMTYDLASAVVRIVNLIGMMLLCHWCCLOFLVPMLO 443
QY 379 GIGTTRW-----VYDGEHNEYLRCYVAVRTLIITIG-GLPEPOTLFEIVFOLLNFFSGV 431
DB 444 DFPDDCWVSINNNVNSWKGYSYALFKAMSHMLCIGYGRQAPVGMDSVWLTWLSMIVGA 503

```

QY 432 FVSSLIQMRDVIGATANQNYFRACMDTDTAYMNNYSIPKLVQRVTVWYVWDSOR 491  
DB 504 TCYAMFIGHATALLQSLDSSRRQYQYKQVEQYMSFHKLPPDTRQRIHDYVYEHRYQG-K 562  
QY 492 MDESDLLKTLPTTQALALADVNF--IISKVDLFGCDTQMIYDMLRLKSLVLYLPG 548  
DB 563 MDESDILGELSEPLREEI---INFNCRKLVASMPLFANADPNFVSMITKLURFEVFPQ 619  
QY 549 DFVCKKGEIGKEMYIIKHGEVOVL-GGPDGTQVLTQKAGSVFGEISLLAAGGNNRRRTAN 607  
DB 620 DVIIEGTIGKMYFIQHGVSVLTKGNKTK----LADGSYFGEICLLTRG---RRTAS 672  
QY 608 VVAHGFANLLTDKKTQLBILVHYDPDSERILMKKA 642  
DB 673 VRADTYCRLYSLSDVNFNEVLEBP-----MMRRA 702

## RESULT 14

US-08-997-685A-2  
; Sequence 2, Application US/08997685A  
; Patent No. 6551821  
; GENERAL INFORMATION:  
; APPLICANT: The Trustees of Columbia University  
; APPLICANT: Kandel, Eric  
; TITLE OF INVENTION: Brain Cyclic Nucleotide Gated Ion Channel and Uses Thereof  
; FILE REFERENCE: 0575/54806  
; CURRENT APPLICATION NUMBER: US/08/997,685A  
; CURRENT FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2

LENGTH: 910  
TYPE: PRT  
ORGANISM: mouse  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (130)..(148)  
OTHER INFORMATION: S1  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (164)..(185)  
OTHER INFORMATION: S2  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (208)..(229)  
OTHER INFORMATION: S3  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (243)..(271)  
OTHER INFORMATION: S4  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (291)..(313)  
OTHER INFORMATION: S5  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (332)..(358)  
OTHER INFORMATION: P  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (367)..(387)  
OTHER INFORMATION: S6  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (472)..(602)  
OTHER INFORMATION: CNB  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: AAC53518  
DATABASE ENTRY DATE: 1997-12-27  
RELEVANT RESIDUES: (1)..(910)  
US-08-997-685A-2

Query Match

9.7%; Score 412.5; DB 2; Length 910;

Best Local Similarity 20.2%; Pred. No. 2.3e-27;  
Matches 150; Conservative 144; Mismatches 269; Indels 179; Gaps 25;

QY 210 IDSYTD-RLYLLWLLLVLTAYNWNCFPLRLVPPYATADNIHYWLIADIICDIIYLVDM 268  
DB 124 IHPYSDFRPY--WDLIMLIMVGNLVIIPVGITP--FTEOTTTPWIFPNVADTVFLLDL 179  
QY 269 LFIQPLRLQVRG-----GDIIVDSNELRKHYRTSTKFLQDVASIIPIFDICVL 315  
DB 180 I-----MNFRTGTVNEDSSEIILDPKVIKMYLXSW-FVDFISSIPVDYIFLIVEKMD 233  
QY 316 -----FFGNPMFRANRMLKYTSFFEFNHLESIMDKAY-----YRVRTTGY 457  
DB 234 SEVYKTARALRIVREFTKILSLRLRLRLRYIHOWEEIFHMTYDVLASAVVRIENL 293  
QY 360 LLFILHINACVY-----WASNYEGIGTRWVYDGEENYLRCLCYWAVRTLIT 407  
DB 294 MLLCHWDGCLQFLVPLQLDPPDCWVSLNE-----MVNSWGQYSAIYLFKAMSHMLC 147  
QY 408 IG-GLPEPQTLEIIVFQLLNFFSGVVFSSLIQMRDVIGATANQNYFRACMDDTIAYM 466  
DB 348 IGYGAQAPVMSDLWITWLSMIVGATCYAMFVGHATALIQLSDSSRRQYQYKQVEQY 407  
QY 467 NNYSPKLVQRVTVWYVWDSQRMDESLDKTLPTTQALALADVNF--IISKVD 523  
DB 408 SPFKLPADMRQKIHDYVYEHRYQG-KIFDEENILSELNDPLREEI---VNFNCRKL 463  
QY 524 LFKGCDTQMIYDMLRLKSLVLYLPGDFVCKKGEIGKEMYIIKHGEVOVLGGPDGT 583  
DB 464 LFANADPNFVTAMLSKLRFEVFPQDGIIRREGAVKMYFIQHGVAAGVI-----TKSKE 518  
QY 584 LKA--GSVFEISLLAAGGNNRRRTANVVAHGFANLLTDKKTQLBILVHYDPDSER 616  
DB 519 MKLTDGSYFGEICLLTKG---RRTASVRADTYCRLYSLSDVNFNEVLEBPMMRA 575  
QY 637 -----ILMKK-----ARULLKOKAK----- 651  
DB 576 AIDRLDRIGKKNLSILLQKFKDLNTGVFNNOENEILKQIVKHDRMVOAIPPI 635  
QY 652 ---TAEATPRKDL-----ALLFP----- 668  
DB 636 LNCSTSTTPTSRMKTQSPVYVATATSLSHSNLHSPSPSTQTPQPSAILSPSYTTAV 695  
QY 669 KEETPKLFTLLGGTGKASLARLLKLKREAAQ---KKENSEGEGEENEDKUKHNE 725  
DB 696 PQSPLATRTTHYASPTASQLSLMQPQOOLPQSOVQOTOTOTOOOQOQOQOQOQO 700  
QY 726 KQKENEKQKENEKDKGRPEEPKPLDRPECTASPIAVEEPHVSRRVTLPRGTSK 744  
DB 756 QQQQQQQQQQQQQQQQQQQQQQPTFGSSTPKNEVHKSTOALHNTLTKVRLPLSAS 815  
QY 785 -----IISMAPSAEGGEEVLT 801  
DB 816 HEVSTLIS-RPHPTVGESLASI 836

## RESULT 15

US-09-086-436-31  
; Sequence 31, Application US/09086436  
; Patent No. 6703485  
; GENERAL INFORMATION:  
; APPLICANT: Kandel, Eric R.  
; APPLICANT: Santoro, Bina  
; APPLICANT: Bartsch, Dusan  
; APPLICANT: Siegelbaum, Steven  
; APPLICANT: Tibbs, Gareth  
; APPLICANT: Grant, Seth  
; TITLE OF INVENTION: Brain or Heart Cyclic Nucleotide Gated Ion Channel and  
; TITLE OF INVENTION: Uses Thereof  
; FILE REFERENCE: 0575/54806-A  
; CURRENT APPLICATION NUMBER: US/09/086,436  
; CURRENT FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 67



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2005, 19:50:54 ; Search time 166 Seconds  
(without alignments)  
2036.289 Million cell updates/sec

Title: US-09-855-828-1  
Perfect score: 4234  
Sequence: 1 MFKSLTKVKNVKPIGENNEN.....PSAEGGEEVLTIEVKEKAKQ 809

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap:  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap:  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap:  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pap:  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pap:  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4234	100.0	809	3	US-09-855-828-1
2	4228	99.9	809	3	US-09-855-828-13
3	3217	76.0	615	5	US-10-978-282-14
4	1952	46.1	1245	5	US-10-978-282-12
5	1781.5	42.1	652	5	US-10-450-763-57830
6	1103	26.1	1037	6	US-11-097-143-26427
7	1088	25.7	644	5	US-10-450-763-35825
8	1052.5	24.9	821	5	US-10-978-282-19
9	955	22.6	747	5	US-10-450-763-42328
10	906	21.4	237	4	US-10-189-507-11
11	902	21.3	237	4	US-10-189-507-7
12	850	20.1	694	3	US-09-842-758-75
13	850	20.1	694	3	US-09-855-828-14
14	850	20.1	694	4	US-10-345-680-26
15	850	20.1	694	4	US-10-174-333-75
16	850	20.1	694	5	US-10-978-282-11
17	829	19.6	663	5	US-10-029-677-16
18	829	19.6	663	5	US-10-978-282-17
19	825.5	19.5	664	4	US-10-029-677-24
20	820.5	19.4	664	4	US-10-029-677-2
21	820	19.4	732	3	US-09-842-758-73
22	820	19.4	732	4	US-10-029-677-15
23	820	19.4	732	4	US-10-174-333-73
24	820	19.4	732	5	US-10-978-282-16
25	819.5	19.4	664	3	US-09-735-927-2
26	819.5	19.4	664	4	US-10-034-843-2
27	819.5	19.4	664	4	US-10-168-651-7

28	819.5	19.4	664	4	US-10-114-153-18	Sequence 18, Appl
29	819.5	19.4	664	5	US-10-978-282-10	Sequence 10, Appl
30	818.5	19.3	664	3	US-09-735-927-4	Sequence 4, Appl
31	818.5	19.3	690	5	US-10-978-282-9	Sequence 9, Appl
32	817	19.3	690	3	US-09-855-828-15	Sequence 15, Appl
33	805.5	19.0	664	5	US-10-978-282-18	Sequence 18, Appl
34	804	19.0	664	4	US-10-029-677-18	Sequence 18, Appl
35	804	19.0	664	4	US-10-087-217-2	Sequence 2, Appl
36	804	19.0	664	4	US-10-295-573-8	Sequence 8, Appl
37	804	19.0	664	5	US-10-978-282-2	Sequence 2, Appl
38	802	18.9	664	4	US-10-295-573-5	Sequence 5, Appl
39	795	18.8	664	4	US-10-087-217-4	Sequence 4, Appl
40	795	18.8	664	5	US-10-978-282-4	Sequence 4, Appl
41	791	18.7	664	4	US-10-087-217-6	Sequence 6, Appl
42	791	18.7	664	4	US-10-295-573-6	Sequence 6, Appl
43	791	18.7	664	5	US-10-978-282-6	Sequence 6, Appl
44	782	18.5	664	4	US-10-087-217-8	Sequence 8, Appl
45	782	18.5	664	5	US-10-978-282-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1  
US-09-855-828-1  
; Sequence 1, Application US/09855828  
; Publication No. US20040137433A1  
; GENERAL INFORMATION:  
; APPLICANT: Crech, Christopher D.  
; APPLICANT: Jegla, Timothy J.  
; APPLICANT: ICAGEN, Inc.  
; TITLE OF INVENTION: CNG3B: A Novel Cyclic Nucleotide-Gated Cation Channel  
; FILE REFERENCE: 018512-00601005  
; CURRENT APPLICATION NUMBER: US/09/855,828  
; CURRENT FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: US 60/204,445  
; PRIOR FILING DATE: 2000-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 809  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human cyclic nucleotide-gated cation channel (CNG)  
; OTHER INFORMATION: 3B (CNG3B)  
US-09-855-828-1

Query Match	100.0%	Score 4234;	DB 3;	Length 809;
Best Local Similarity	100.0%;	Pred. No. 5.1e-301;		
Matches 809;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MFKSLTKVKNVKPIGENNENEQSSRRNEGSHPSNOSQQTAAQENKGEKSLTKSTPV	60	
Db	1	MFKSLTKVKNVKPIGENNENEQSSRRNEGSHPSNOSQQTAAQENKGEKSLTKSTPV	60	
Qy	61	TSSEPHNTIQQKLSKKNSSGDLTTNPDQNAASPTGTVPEQKMDPGKEGPNQKPPA	120	
Db	61	TSSEPHNTIQQKLSKKNSSGDLTTNPDQNAASPTGTVPEQKMDPGKEGPNQKPPA	120	
Qy	121	APVINEYADQALNLYKRMQRORTALYKKLVEGDLSSPEASPTAKPTAVPPVKESDDKP	180	
Db	121	APVINEYADQALNLYKRMQRORTALYKKLVEGDLSSPEASPTAKPTAVPPVKESDDKP	180	
Qy	181	TEHYRYLLWFKVKMPLTEYLKIKLPNSIDSYTDRLYLWLLVTLAYNWCWFIPLRL	240	
Db	181	TEHYRYLLWFKVKMPLTEYLKIKLPNSIDSYTDRLYLWLLVTLAYNWCWFIPLRL	240	
Qy	241	VFPYQTADNTHYWLIIADIIICDIIYLDMLFIQPRLOFVRGCDIIVDSNELRKHYRTSTKF	300	
Db	241	VFPYQTADNTHYWLIIADIIICDIIYLDMLFIQPRLOFVRGCDIIVDSNELRKHYRTSTKF	300	
Qy	301	QLDVASIIIPDICYLFFGFNPMFRANMLKTYTSFFEFNHHLESIMDKAYIYRVIRTTGYL	360	

Db 301 QLDVASIIPFDICYLFFGFGNPMFRANRMLKYTSFFEFNHHLESIMDKAYIRVIRTTGYL 360  
QY 361 LFILHINACVYVWASNYEGITTRWVYDGEHNEYLRCYVWAVRTLTITIGGLPEPOTLFEI 420  
Db 361 LFILHINACVYVWASNYEGITTRWVYDGEHNEYLRCYVWAVRTLTITIGGLPEPOTLFEI 420  
QY 421 VFQQLNFFSGVVFSSSLIGQMRDVI GAATANQNYFRACMDDTIAYMNNYSIPKL VQKRV 480  
Db 421 VFQQLNFFSGVVFSSSLIGQMRDVI GAATANQNYFRACMDDTIAYMNNYSIPKL VQKRV 480  
QY 481 TWYEYTWDSQRMLESLLKTLPTTVQLALADVNFSIISKVDLFGKCDTQMIYDMLLRL 540  
Db 481 TWYEYTWDSQRMLESLLKTLPTTVQLALADVNFSIISKVDLFGKCDTQMIYDMLLRL 540  
QY 541 KSVLYLPDGFVCKKGIGKEMVIIKHGEVQVGLGGPDGTVLTAKAGSVFGEISLLAAGG 600  
Db 541 KSVLYLPDGFVCKKGIGKEMVIIKHGEVQVGLGGPDGTVLTAKAGSVFGEISLLAAGG 600  
QY 601 GNRRTANVVAHGFANLLTLDKKTLOEILVHYPDSEIRILMKKARVLLKQAKTAEATPPRK 660  
Db 601 GNRRTANVVAHGFANLLTLDKKTLOEILVHYPDSEIRILMKKARVLLKQAKTAEATPPRK 660  
QY 661 DLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREQAQKENSEGEGEGEKENEDKQ 720  
Db 661 DLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREQAQKENSEGEGEGEKENEDKQ 720  
QY 721 KENEDKQENEDKGENEDKDKGREPEEKPLDRPECTASPIAVEEPHVSVRTVLPGRGTS 780  
Db 721 KENEDKQENEDKGENEDKDKGREPEEKPLDRPECTASPIAVEEPHVSVRTVLPGRGTS 780  
QY 781 RQSLIISMAPSAEGGEEVLTIEVKEKAKQ 809  
Db 781 RQSLIISMAPSAEGGEEVLTIEVKEKAKQ 809

RESULT 2

US-09-855-828-13  
; Sequence 13, Application US/09855828  
; Publication No. US20040137433A1  
; GENERAL INFORMATION:  
; APPLICANT: Creech, Christopher D.  
; APPLICANT: Jegia, Timothy J.  
; APPLICANT: ICAGEN, Inc.  
; TITLE OF INVENTION: CNG3B: A Novel Cyclic Nucleotide-Gated Cation Channel  
; FILE REFERENCE: 018512-006010US  
; CURRENT APPLICATION NUMBER: US/09/855, 828  
; CURRENT FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: US 60/204,445  
; PRIOR FILING DATE: 2000-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 809  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: CNG3B  
US-09-855-828-13

Query Match 99.9%; Score 4228; DB 3; Length 809;  
Best Local Similarity 99.9%; Pred. No. 1.4e-300;  
Matches 808; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFKSLTKVNVKPIGENNENESSRRNEGSHPSNOSQOTTAQENKGEESLKTKSTPV 60  
Db 1 MFKSLTKVNVKPIGENNENESSRRNEGSHPSNOSQOTTAQENKGEESLKTKSTPV 60  
QY 61 TSEEPHTNIQDKLSKKNSSGDLTTNPDQNAAEPTGTVEQKEMDPGKGNPSQNKPPA 120  
Db 61 TSEEPHTNIQDKLSKKNSSGDLTTNPDQNAAEPTGTVEQKEMDPGKGNPSQNKPPA 120  
QY 121 APVINEYADAQLHNLVKRMQROTALYKKKLVEGDLSSPEASPTAKPTAVPPVKESDDKP 180

Db 121 APVINEYADAQLHNLVKRMQROTALYKKKLVEGDLSSPEASPTAKPTAVPPVKESDDKP 180  
QY 181 TEHYVRLWLFVKWKMPLEYLKRILKLPNSIDSYTDRLVLLWLLVTLAYNNCWFIPURL 240  
Db 181 TEHYVRLWLFVKWKMPLEYLKRILKLPNSIDSYTDRLVLLWLLVTLAYNNCWFIPURL 240  
QY 241 VFPYQTADNIHWYLIADIICDIILYDMLFQPRLOQFVRGGDIIIVDSNELRKHRYTSTKF 300  
Db 241 VFPYQTADNIHWYLIADIICDIILYDMLFQPRLOQFVRGGDIIIVDSNELRKHRYTSTKF 300  
QY 301 QLDVASIIPFDICYLFFGFGNPMFRANRMLKYTSFFEFNHHLESIMDKAYIRVIRTTGYL 360  
Db 301 QLDVASIIPFDICYLFFGFGNPMFRANRMLKYTSFFEFNHHLESIMDKAYIRVIRTTGYL 360  
QY 361 LFILHINACVYVWASNYEGITTRWVYDGEHNEYLRCYVWAVRTLTITIGGLPEPOTLFEI 420  
Db 361 LFILHINACVYVWASNYEGITTRWVYDGEHNEYLRCYVWAVRTLTITIGGLPEPOTLFEI 420  
QY 421 VFQQLNFFSGVVFSSSLIGQMRDVI GAATANQNYFRACMDDTIAYMNNYSIPKL VQKRV 480  
Db 421 VFQQLNFFSGVVFSSSLIGQMRDVI GAATANQNYFRACMDDTIAYMNNYSIPKL VQKRV 480  
QY 481 TWYEYTWDSQRMLESLLKTLPTTVQLALADVNFSIISKVDLFGKCDTQMIYDMLLRL 540  
Db 481 TWYEYTWDSQRMLESLLKTLPTTVQLALADVNFSIISKVDLFGKCDTQMIYDMLLRL 540  
QY 541 KSVLYLPDGFVCKKGIGKEMVIIKHGEVQVGLGGPDGTVLTAKAGSVFGEISLLAAGG 600  
Db 541 KSVLYLPDGFVCKKGIGKEMVIIKHGEVQVGLGGPDGTVLTAKAGSVFGEISLLAAGG 600  
QY 601 GNRRTANVVAHGFANLLTLDKKTLOEILVHYPDSEIRILMKKARVLLKQAKTAEATPPRK 660  
Db 601 GNRRTANVVAHGFANLLTLDKKTLOEILVHYPDSEIRILMKKARVLLKQAKTAEATPPRK 660  
QY 661 DLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREQAQKENSEGEGEGEKENEDKQ 720  
Db 661 DLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREQAQKENSEGEGEGEKENEDKQ 720  
QY 721 KENEDKQENEDKGENEDKDKGREPEEKPLDRPECTASPIAVEEPHVSVRTVLPGRGTS 780  
Db 721 KENEDKQENEDKGENEDKDKGREPEEKPLDRPECTASPIAVEEPHVSVRTVLPGRGTS 780  
QY 781 RQSLIISMAPSAEGGEEVLTIEVKEKAKQ 809  
Db 781 RQSLIISMAPSAEGGEEVLTIEVKEKAKQ 809

RESULT 3

US-10-978-282-14  
; Sequence 14, Application US/10978282  
; Publication No. US20050221426A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, Yong  
; APPLICANT: CAO, Liang  
; APPLICANT: LU, Jianming  
; APPLICANT: LLORENTE, Isabel  
; TITLE OF INVENTION: Novel Cell-Based Assays Employing Voltage and Calcium Dyns  
; FILE REFERENCE: 062614-5002-US  
; CURRENT APPLICATION NUMBER: US/10/978,282  
; CURRENT FILING DATE: 2004-11-01  
; PRIOR APPLICATION NUMBER: US 60/589,012  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: US 60/515,442  
; PRIOR FILING DATE: 2003-10-30  
; PRIOR APPLICATION NUMBER: US 10/087,217  
; PRIOR FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: US 60/330,663  
; PRIOR FILING DATE: 2001-10-26  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 615

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(615)
; OTHER INFORMATION: CNGB3
US-10-978-282-14

Query Match 76.0%; Score 3217; DB 5; Length 615;
Best Local Similarity 100.0%; Pred. No. 1.1e-226;
Matches 615; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 195 MPEYLKRIKIPNSIDSTDRLLYLWLLVLTAYNNWCWFILRLVFPYQTADNIHYWL 254
Db 1 MPEYLKRIKIPNSIDSTDRLLYLWLLVLTAYNNWCWFILRLVFPYQTADNIHYWL 60

Qy 255 IADIICDIYLYDMLFIQRLQFVRGDIIVDSNELRKHRTSTKFLQDVASIIIPEDICY 314
Db 61 IADIICDIYLYDMLFIQRLQFVRGDIIVDSNELRKHRTSTKFLQDVASIIIPEDICY 120

Qy 315 LFFGFNPMFRANRMLKYTSFFFNHLESIMDKAYIYRVIRTTGYLLFILHINACYYYWA 374
Db 121 LFFGFNPMFRANRMLKYTSFFFNHLESIMDKAYIYRVIRTTGYLLFILHINACYYYWA 180

Qy 375 SNYEGIGTRWYVDGEGNEYLRCYYWAVRTLITIGGLPEPQTLFEIVFOLLNPFSGVFVF 434
Db 181 SNYEGIGTRWYVDGEGNEYLRCYYWAVRTLITIGGLPEPQTLFEIVFOLLNPFSGVFVF 240

Qy 435 SSLIGQMRDVIGAAATANQYFRACMDTTIAYMNNYSIPKLQKRVRTWYETWDSORMLD 494
Db 241 SSLIGQMRDVIGAAATANQYFRACMDTTIAYMNNYSIPKLQKRVRTWYETWDSORMLD 300

Qy 495 ESDLLKTLPTTVQALALDYNFISISKVDLFGKCDTQMIYDMLRLKSLVLYLPGDFVCKK 554
Db 301 ESDLLKTLPTTVQALALDYNFISISKVDLFGKCDTQMIYDMLRLKSLVLYLPGDFVCKK 360

Qy 555 GETGKEMYYIKHGEVQVLGPDGTVLTKAGSVFGEISLLAAGGNNRTANVVAHGFA 614
Db 361 GETGKEMYYIKHGEVQVLGPDGTVLTKAGSVFGEISLLAAGGNNRTANVVAHGFA 420

Qy 615 NLTLDKKTLQEILVHPDSEIRILMKKARVLLKQAKTAEATPPKDLALLFPKSETPK 674
Db 421 NLTLDKKTLQEILVHPDSEIRILMKKARVLLKQAKTAEATPPKDLALLFPKSETPK 480

Qy 675 LFKTLGGTGKASLARLLKLRQAAQKENSGBEGEKENEDKOKENEDK 734
Db 481 LFKTLGGTGKASLARLLKLRQAAQKENSGBEGEKENEDKOKENEDK 540

Qy 735 KENEDKDKGREBEKPLDRPECTASPIAVBEEPHSVRRTVLPRTGTSRQSIIISMPSAEG 794
Db 541 KENEDKDKGREBEKPLDRPECTASPIAVBEEPHSVRRTVLPRTGTSRQSIIISMPSAEG 600

Qy 795 GEEVLTIIEVKEKAKQ 809
Db 601 GEEVLTIIEVKEKAKQ 615

RESULT 4
US-10-978-282-12
; Sequence 12, Application US/10978282
; Publication No. US20050221426A1
; GENERAL INFORMATION:
; APPLICANT: YAO, Yong
; APPLICANT: CAO, Liang
; APPLICANT: LU, Jiaming
; APPLICANT: LORENTE, Isabel
; TITLE OF INVENTION: Novel Cell-Based Assays Employing Voltage and Calcium Dyes
; FILE REFERENCE: 062614-5002-US
; CURRENT APPLICATION NUMBER: US/10/978,282
; CURRENT FILING DATE: 2004-11-01
; PRIOR APPLICATION NUMBER: US 60/589,012
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: US 60/515,442
```

```
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US 10/087,217
; PRIOR FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/330,663
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1245
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(1245)
; OTHER INFORMATION: CNGB1
US-10-978-282-12

Query Match 46.1%; Score 1952; DB 5; Length 1245;
Best Local Similarity 46.4%; Pred. No. 1.3e-133;
Matches 405; Conservative 121; Mismatches 237; Indels 110; Gaps 18;

Qy 14 IGENNENEQSSRRNEEGSHPSNQSOOTTAQE--ENKGE---EKSUKTKTPTVTSERPHTN 69
Db 403 VGEAEKAEKAEKAEAEAEAEAEAEKQDWAETKEEPEAEAEAEASSGVPATKQHPEVQ 462

Qy 69 IQDK-----LSKNSSGDLTTNPDQNAAEPTGTVPQEKMDPKCEGPN----- 112
Db 463 VEDTDADSCPLMAEENPPS--TVLPPSPAKSDTLIVSSASGTHRKKLUPSEDEAEELK 520

Qy 113 --SPQKRP-----PAAP-----VINEYADAOLHNLVKMRORTALY 146
Db 521 ALSPAESPVVAMSDPTTPKDTGQDRAASTASTNSAIIND----RLQELVKLFKERTKV 576

Qy 147 KKKLVGEGDLS----SPEASPTAKPTAVPPVKESDDKPT--EHYRLLLPFKVKKMPLETEY 200
Db 577 KEKLDIPDVTSDSESPKSPAKKAPAPDTPKPAEAEPEVEEHYCDMLCKCKFKHRPWKKY 636

Qy 201 LKRIKLPNSIDSYDRLLYLWLLVLTAYNNWCWFILRLVFPYQTADNIHYWLIADIIC 260
Db 637 ---QFQSIDPLTNLNYVLLFVVMWNNWCWLIPIVRWAPFYOTPDNIHHMLLMDYLC 692

Qy 261 DIILYDMLFIQRLQFVRGDIIVDSNELRKHRTSTKFLQDVASIIIPFIDICLFFGFN 320
Db 693 DLIFLFDITVQTRLOFVRGDDIITDKDMRNLYKSRFRKMDLSSLPLDLYLKGVN 752

Qy 321 PMFRANRMLKYTSFFFNHLESIMDKAYIYRVIRTTGYLLFILHINACYYYWASNYEGI 380
Db 753 PLLRLPRLKYNWAFEFNSRLESILSKAYVYRVIRTTAVLLYSLHNSCLYYWASAYOGL 812

Qy 381 GTTRWYVDGEGNEYLRCYYWAVRTLITIGGLPEPQTLFEIVFOLLNPFSGVFSSLIQ 440
Db 813 GSTHWYVDGVGNSYIRCYFAVKTLITIGGLPQPKTLFEIVFOLLNPFSGVFASFVNIQ 872

Qy 441 MRDVIGAAATANQYFRACMDTTIAYMNNYSIPKLQKRVRTWYETWDSORMLDESLLK 500
Db 873 MRDVIGAAATAGQTYRSCMDSTVKYMNFKIPKSVQNRVKTWYETWHSOGLMDESELMV 932

Qy 501 TLPTTVQALALADYNFISISKVDLFGKCDTQMIYDMLRLKSLVLYLPGDFVCKKEIGKE 560
Db 933 QLPDKMRDLADADYNVIVSKVALFOGCDRQMIQFDMKLRLSVVLYLPNDVYCKKKEIGRE 992

Qy 561 MYIIKHGEVQVLGPDGTVLTKAGSVFGEISLLAAGGNNRTANVVAHGFAANLLTLD 620
Db 993 MYIIQAGQVQLGPDGKSVLTKAGSVFGEISLLAAGGNNRTANVVAHGFTNLFI 1052

Qy 621 KKTQLQILVHPDSEIRILMKKARVLLKQAKTAEATPPKDLALLFPKKEETPKLFTLL 680
Db 1053 KQDLNAILVHPESQKLLRKKARMLRSNNK-----PKEEKSVLIIPPRAGTPKLFAAL 1107

Qy 681 GGTGKAS-----LARLLKLRQAAQKENSGBEGEKENEDKOKENEDK 728
Db 1108 AMTKMGKGKAKGKGLAHLRLKELAALEAAKQDEL----VEQAKSSODVKEGESAAP 1164
```

```
QY 729 ENEDKGKE-NEDKDKGREPERK-----LDRPECTASPIAVEREPEHSVRTVLP 776
DQ 1165 DQTHPKAATDPPAPRTPEPPGPPSPPPASLGRPEGBEGA-EPEEHSVR----- 1218
QY 777 RGTSRQSLIISMAPSAEGGEEVLTIEVKEKAKQ 809
DQ 1219 -----ICWSPGPEGEQILSVKMPPEERBE 1242

RESULT 5
US-10-450-763-57830
; Sequence 57830, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CI23/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 57830
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (387)..(411)
; OTHER INFORMATION: Cyclic nucleotide-binding domain proteins domain identified
; OTHER INFORMATION: by eMATRIX, accession number BL00888B, p-value=9.143e-16, raw score
; OTHER INFORMATION: of 14.79
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (133)..(349)
; OTHER INFORMATION: Transmembrane region cyclic Nucleotide G domain identified by
; OTHER INFORMATION: Pfam, accession name CNG_membrane, E-values=2.4e-92, Pfam score of
; OTHER INFORMATION: 320.2
US-10-450-763-57830

Query Match 42.1%; Score 1781.5; DB 5; Length 652;
Best Local Similarity 55.3%; Pred. No. 1.6e-121; Indels 47; Gaps 7;
Matches 343; Conservative 87; Mismatches 143;

QY 214 TDRLLMLLLVTLAYNNWCNFIPLRLFPYQADNIHWHLIADICDIYLYDMLFIQP 273
DQ 53 TDLMTVLMFFVMAWNNCWLPVRWAPPYQTPDNIHWHLLMDVLCDDLIYFLDITVQT 112
QY 274 RLQFVRGDIIVDSNELRKHRTSTKFDLVAIIPDFICYLFFGNPMFRANRLKYTS 333
DQ 113 RLQFVRGDIITDKDMRNLYKSRFRKMDLSLLPLDFLYLKVGVPNLLRPLRCLKYMA 172
QY 334 PFEFNHLESIMDKAVIVRVITGTGLFILHINACVYVWASNYEGIGTTRVWYDEGNE 393
DQ 173 PFEFNHLESILSKAYVVRVITAYLLSLHNSCLYVWASAYQGLGTHWYDGVGNS 232
QY 394 YLRCYVAVRTLITIGLPEPOTLFEIVPQLNFFSGVVFVSSLIGQMRDVIIGAATANQN 453
DQ 233 YIRCYFAVAKTLITIGLPPDKTLFEIVPQLNFFSGVVFVSSLIGQMRDVIIGAATANQN 292
QY 454 YFRACMDTIAWNNYSIPKLVOKVRVYETWDSQRMDSLLKTLPTTVQLALAI 513
DQ 293 YRSCMDSTVKMNFYKIPKSVQNVKTYETWHSQGLMDESLMVQLPDKRDLAI 352
QY 514 VNFSIISKVDLPFGCDTOMIYDMLRLKLSVLYPGDFVCKGKEIGKEMVIIKHGVOVLG 573
DQ 353 VNYNIVSKVALFQGDRCQMFDMRLKRLRSVVLYPNDYVCKGKEIGREMYIIQAGQVQLG 412
```

```
QY 574 GPDGTYLVLTIKAGSVFGEISLLAAGGNNRRTANVVAHGFANLLTLDKKTLOEILVHYPD 633
DQ 413 GPDGTYLVLTIKAGSVFGEISLLAAGGNNRRTANVVAHGFANLLTLDKKTLOEILVHYPD 472
QY 634 SERILMKKARVLLKOKAKTAEATPRKDLALLFPPEETPKLFTLLLGTTGKAS----- 687
DQ 473 SQKLLRKARRMLRSNNK-----PKEKSVLILPRACTPKLFNAALAMTKMGMGKGAGK 527
QY 688 -----LARLLKREOAAOKKENSEGEGEKENEDKOKENEDKKE-NEDK 740
DQ 528 GKLAHLRLARLAKLALEAAAKHEEL-----VEQAKSODVKGEESAAPQOHTPKKAAATDP 584
QY 741 DKGREPEEKPLDRPECTASPIAV-----EEPHSVRRTVLPRTGTSRQSLIISMA 789
DQ 585 PAPRTPEPP-GSPSSPPPASLGSGEGEESGPAEPBHSVR----- 849
QY 790 PSAEGGEEVLTIEVKEKAKQ 809
DQ 630 PGPEPEGEQILSVKMPPEEREE 649

RESULT 6
US-11-097-143-26427
; Sequence 26427, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26427
; LENGTH: 1037
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-26427

Query Match 26.1%; Score 1103; DB 6; Length 1037;
Best Local Similarity 33.4%; Pred. No. 1.5e-71;
Matches 243; Conservative 153; Mismatches 255; Indels 76; Gaps 13;
```

```
QY 16 ENNEEQSSRRNEGSHPSNQSQTTAQ-----EENKGEKSLKTKSTPTVTSSEPHNTI 69
DQ 276 EEEEEKSSPLHOVESQDVDEQDVQICYNESPELONEDRNETORTPSINESEIVAV 335
QY 70 QDKLSKKNSSGDLTTNPDQNAAEFTGTVPQEKMDPGKEGPNQPNKPAAPVINEYAD 129
DQ 336 DEPTDMNSVD---HRKFPSSA---GSLDSQSQ-----QFLR 366
QY 130 AOLHNLVKMRQRTALYKKLVGDLSSPEASPOKAPTAVPPVKESDDKTEHYRLLW 189
DQ 367 DQVRHLVRRRTARANKVKSRI---ELPTPTSSSTVSSPPTKSLHSP-QH----- 416
```

```
QY 190 PKYKMPLETKRIKL-----PNSIDSYTD---RLYLWLLLVTLA 228
Db 417 -KVLVPAGQSPHGRGLFEADTPRSNVMLCSSLGANDERTLDPQGIYISWLCVSL 475
QY 229 YNWNCFIPILRLVPEYQADNIHYMLADIIDIIYLYDMLFIQPLQFVRGDDIIVDSN 288
Db 476 FLYNAVWIFILRASFPQTKENTINWILACDFADIIYLLDVFPFKHVMYLFSGFWKKN 535
QY 289 ELRKHYRTSTKFOLDVASIIPEDICYLFFGPMF--RANRMLKYTSFFPFNFHLESIMDK 347
Db 536 LTRKNTYRKLOPKDLALLPLELLYFKLTQAVWLRFPRFKIQSFWEVFRLLDRVISS 595
QY 348 AYIYRVIRTTGYLLFILHINACVYVWASNYBEGITTRWYDGBGNYLRCYVWVRLTIT 407
Db 596 PHFVRVAKTLTYMLYMIHTAALYAYSDYQGLQNRWVFSKGHPYVRCFAPATKTATS 655
QY 408 IGGLEPQTLPEIVFOLLNFFSIVFSSILGOMRDVIGANTANQNYFRACDDDTIAYVN 467
Db 656 IGKPKPERQGEYVFMVAWLVGFVALLIGQIRDIISTATNKGHEYQLEDETLEYMR 715
QY 468 NYSIPKLVQKRVRTYETWDSORMLDESLLKTLPTTVQLALAIIDVNFISIISKVDFKG 527
Db 716 RLNLSEVQSRVQWFOFTWEQRTLDSENLIDALINLKTDAISVHIOTLSKVQLFAD 775
QY 528 CDTQMIYDMLLRKSVLYLPDGPVCKKGEIGKEMYYIHKGEVQVLGGPDGTRKVLTKAG 587
Db 776 CBEALLRDLVLKRAVTFLPDGPVCKRGEVREMYIVKLQGVQMGSPSDVVVLTATLTEG 835
QY 588 SVFGEISLLAAGGNRTANVAHGFANLLTLDKKTLOEILVHYPDSEIRILMKKARVLL- 646
Db 836 SVFGEISLLGADRRRTADVRSGYSLNFVLSKSDNLNEVIAYYTAQAILKGRQOLMR 895
QY 647 KOKAKTAEATPPRKDAL-----LFPPK--EETPKLFTLLGGTGKASLARLLKLKREQA 699
Db 896 KNAREEERERARSALQADVIGNPKTETAPKLLQTVIQALPESPASVAVLITRSGKR 955
QY 700 AQKENS 706
Db 956 MRRKQS 962

RESULT 7
US-10-450-763-35825
; Sequence 35825, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 35825
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (58)..(105)
; OTHER INFORMATION: kw TRANSCRIPTASE REVERSE II ORF2 domain identified by
; OTHER INFORMATION: eMATRIX, accession number DM01354V, p-value=1.000e-40, raw score
; OTHER INFORMATION: 12.97
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(644)
```

; OTHER INFORMATION: Xaa = X or \* as defined in Table 2  
US-10-450-763-35825

Query Match 25.7%; Score 1088; DB 5; Length 644;  
Best Local Similarity 100.0%; Pred. No. 9.9e-71;  
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 594 SLLAAGGNRRRTANVAHGFANLLTLDKKTLOEILVHYPDSEIRILMKKARVLLKOKAKTA 653  
Db 429 SLLAAGGNRRRTANVAHGFANLLTLDKKTLOEILVHYPDSEIRILMKKARVLLKOKAKTA 488  
QY 654 EATPPRKDLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREQAQAOKKENSEGEEBEG 713  
Db 489 EATPPRKDLALLFPKKEETPKLFTLLGGTGKASLARLLKLKREQAQAOKKENSEGEEBEG 548  
QY 714 KENEDKQKENEKQKENEKQKENEKQKREPEKPLDRPECTASPIAVEEPHSHVRRT 773  
Db 549 KENEDKQKENEKQKENEKQKENEKQKREPEKPLDRPECTASPIAVEEPHSHVRRT 608  
QY 774 VLPRTSROSLLISMAPSASGEEVLTIEVKEKAKQ 809  
Db 609 VLPRTSROSLLISMAPSASGEEVLTIEVKEKAKQ 644

RESULT 8

US-10-978-282-19  
; Sequence 19, Application US/10978282  
; Publication No. US20050221426A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, Yong  
; APPLICANT: CAO, Liang  
; APPLICANT: LU, Jianming  
; APPLICANT: LLORENTE, Isabel  
; TITLE OF INVENTION: Novel Cell-Based Assays Employing Voltage and Calcium Dyes  
; FILE REFERENCE: 062614-5002-US  
; CURRENT APPLICATION NUMBER: US/10/978,282  
; CURRENT FILING DATE: 2004-11-01  
; PRIOR APPLICATION NUMBER: US 60/589,012  
; PRIOR FILING DATE: 2004-07-20  
; PRIOR APPLICATION NUMBER: US 60/515,442  
; PRIOR FILING DATE: 2003-10-30  
; PRIOR APPLICATION NUMBER: US 10/087,217  
; PRIOR FILING DATE: 2002-03-04  
; PRIOR APPLICATION NUMBER: US 60/330,663  
; PRIOR FILING DATE: 2001-10-26  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 821  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: human CNGA and CNGB consensus sequence  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(821)  
; OTHER INFORMATION: Xaa = any naturally occurring amino acid  
US-10-978-282-19

Query Match 24.9%; Score 1052.5; DB 5; Length 821;  
Best Local Similarity 38.8%; Pred. No. 5.6e-68;  
Matches 227; Conservative 97; Mismatches 205; Indels 59; Gaps 6;

QY 205 KLPNSIDSYTDRLYLLWLLVTLAYNNWNCWFILRLVFPYQYADNIHYMLADIICDIY 264  
Db 256 KLLXVIDPSCDXLYLLWLFXIALPVNNWCLLVARACFPDLQSDYLYHVLVDVSDVY 315  
QY 265 LYDMLFIQPLQFVRGDDIIVDSNEIRKHYRTSTKTFOLDVASIIPDICYLFFGPN-PMF 323  
Db 316 LLDMXFVRTGTGEQGLLVVDNKLNNYKTTLQFKLDVASLIPTDLYLVKVGXNYPEV 375  
QY 324 RANRMLKYTSFFPFNFHLESIMDKAYIYRVIRTTGYLLFILHINACVYVWASNYEGIGTT 383



Db 1 LFEIVFOLLNFTYGVAFSPVMIQMRDVVVGAAATAGCTYRSCMDSTVKVMNFKYKIPKSVQ 60  
QY 477 KRVRTWYETWDSQRLMDESLLKTLPTTVQLALADIVNFSISKVDLFGCDTQMIYDM 536  
Db 61 NRKVTWYETWHSQGLMDESELVQLPDKMRDLDAIDVNYNIVSKVALFQGCDCRQWIFDM 120  
QY 537 LLRLKSVLYLPDGFVCKKGEIGEMKYYIKHGEVQVLGGPDGTVKVLVTLKAGSVFGEISLL 596  
Db 121 LKRLSVLYLPNDYVCKKGEIGREMYIIQAGQVQLGGPDGKSVLTVTLKAGSVFGEISLL 180  
QY 597 AAGGNRRRTANVAVHGFANLLTLDKKTQELVHYVDPSEIRILMKKARVLLKQAK 651  
Db 181 AVGGNRRRTANVAVHGFNTLFLDKKDLNEILLVHYVESQKLKRRKARRMLRNNK 235

RESULT 11  
US-10-189-507-7  
; Sequence 7, Application US/10189507  
; Publication No. US20030228633A1  
; GENERAL INFORMATION:  
; APPLICANT: ZOLLER, MARK  
; APPLICANT: XU, HONG  
; APPLICANT: STASZEWSKI, LENA  
; APPLICANT: MOYER, BRYAN  
; APPLICANT: PROMIN, ALEX  
; APPLICANT: ADLER, JON ELLIOT  
; APPLICANT: SERVANT, GUY  
; APPLICANT: CALLAMARAS, NICHOLAS  
; TITLE OF INVENTION: EXPRESSION OF FUNCTIONAL HUMAN OLFACTORY CYCLIC  
; TITLE OF INVENTION: NUCLEOTIDE GATED (CNG) CHANNEL IN RECOMBINANT HOST  
; TITLE OF INVENTION: CELLS AND USE THEREOF IN CELL BASED ASSAYS TO IDENTIFY  
; TITLE OF INVENTION: SHELLED MODULATORS  
; FILE REFERENCE: 078003-0291567  
; CURRENT APPLICATION NUMBER: US/10/189,507  
; CURRENT FILING DATE: 2003-02-12  
; PRIOR APPLICATION NUMBER: 60/303,140  
; PRIOR FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/337,154  
; PRIOR FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO: 7  
; LENGTH: 237  
; TYPE: 'PRT'  
; ORGANISM: Rattus sp.  
US-10-189-507-7

Query Match 21.3%; Score 902; DB 4; Length 237;  
Best Local Similarity 72.3%; Pred. No. 1e-57;  
Matches 170; Conservative 30; Mismatches 35; Indels 0; Gaps 0;  
QY 417 LPEIVFOLLNFTSGVFSGLIQMRDVICAATANQYFRACMDDTIAYMNNYSIPLVQ 476  
Db 1 LFEIVFOLLNFTYGVAFSPVMIQMRDVVVGAAATAGCTYRSCMDSTVKVMNFKYKIPKSVQ 60  
QY 477 KRVRTWYETWDSQRLMDESLLKTLPTTVQLALADIVNFSISKVDLFGCDTQMIYDM 536  
Db 61 NRKVTWYETWHSQGLMDESELVQLPDKMRDLDAIDVNYNIVSKVALFQGCDCRQWIFDM 120  
QY 537 LLRLKSVLYLPDGFVCKKGEIGEMKYYIKHGEVQVLGGPDGTVKVLVTLKAGSVFGEISLL 596  
Db 121 LKRLSVLYLPNDYVCKKGEIGREMYIIQAGQVQLGGPDGKSVLTVTLKAGSVFGEISLL 180  
QY 597 AAGGNRRRTANVAVHGFANLLTLDKKTQELVHYVDPSEIRILMKKARVLLKQAK 651  
Db 181 AVGGNRRRTANVAVHGFNTLFLDKKDLNEILLVHYVESQKLKRRKARRMLRNNK 235

RESULT 12  
US-09-842-758-75  
; Sequence 75, Application US/09842758  
; Publication No. US20030083244A1  
; GENERAL INFORMATION:

; APPLICANT: Vernet, Corine A. M.  
; APPLICANT: Fernandes, Elma R  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Shinkets, Richard A  
; APPLICANT: Maiyankar, Uriel M  
; APPLICANT: Boldog, Ferenc L  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Tchervnev, Velizar T  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Rastelli, Luca  
; APPLICANT: MacDougall, John R  
; APPLICANT: Taupier, Raymond J  
; APPLICANT: Grosse, William M  
; APPLICANT: Edward, Szekeres S  
; APPLICANT: Alsobrook II, John P  
; TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-783  
; CURRENT APPLICATION NUMBER: US/09/842,758  
; CURRENT FILING DATE: 2001-04-25  
; PRIOR APPLICATION NUMBER: 60/200,158  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: 60/200,613  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,780  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/201,006  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/201,007  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/201,236  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/201,238  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/201,186  
; PRIOR FILING DATE: 2000-05-02  
; PRIOR APPLICATION NUMBER: 60/201,474  
; PRIOR FILING DATE: 2000-05-03  
; PRIOR APPLICATION NUMBER: 60/201,508  
; PRIOR FILING DATE: 2000-05-03  
; PRIOR APPLICATION NUMBER: 60/220,591  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: 60/232,678  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: 60/263,217  
; PRIOR FILING DATE: 2001-01-22  
; PRIOR APPLICATION NUMBER: 60/265,160  
; NUMBER OF SEQ ID NOS: 113  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 75  
; LENGTH: 694  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-842-758-75

Query Match 20.1%; Score 850; DB 3; Length 694;  
Best Local Similarity 30.0%; Pred. No. 3e-53;  
Matches 225; Conservative 149; Mismatches 269; Indels 106; Gaps 22;  
QY 27 NEEGSHPSNQSQTTAQEENKGEESKLTSTPTVSEEPHTNIQDKLSKKNSSGDLTTNP 86  
Db 5 NTQYSHPS-----RTHLKVK-----TSDRLNRAENGLSRAHSSSEETS-- 43  
QY 87 DPONAAEPTGTVPQKEMDPGKGNPSQNKPPAAPVINEYADAOLHNLVKRNRQTALY 146  
Db 44 ---SLVLP-GIANETRGADSGGSGFTGQ-----GIARLSRLIFLLRRWAARH 87

QY 147 KKLVGDLSSP-----EASQTAKTAVPPVKESDDKPTBHYVRLMLFKVKKMP 197  
 DB 88 VHQDQGPDPDFRFGAELKEVSSQESNAQNVGSEPADRG-----RSAW-----PL 136  
 QY 198 -----TEYLKRIKLPNSI-DSYTDRLVLLMLLVTLAYNNCWFPIRLRVPVOT 246  
 DB 137 AKCNTNTSNTTEEEKTKKDAIIVDPSSNLYRMLTAIALPVFNWYLLICRACFDELQ 196  
 QY 247 ADNIHWLIADIICDIYLYDMLFIQPRLOFVRGGDIIVDSNELRKHYRTSTKFFOLDVAS 306  
 DB 197 SEYLMMLVLDYSADVLYLDVL-VRAATGFLQGLMVSDTNRLWQHVKTTTQFKLDVLS 255  
 QY 307 LIIPDICYLFQFN-PMFRANRMLKTSFFEFNHLESIMDKAYIYRVIRTTGYLLFIH 365  
 DB 256 LVPTDLAYLVKGTNYPEVRFNRLKFSRLEFFDRTETRTNYPNMFRLGNLVLYLIIH 315  
 QY 366 INACVYVWASNYEGIGTRWVYDG-----EGNEYLRCYVAVRTLTITIGLPEPOTL 417  
 DB 316 WNACIYFAISKFIQGTDSWVYPNISIPHGRLSRKIYSLWSTLTTLTTIGETPPVKD 375  
 QY 418 FEIVFOLLNFFSGVVFSSLIQMRDVIQAATANQYFRACDDTIAYNNYSIPKLQVK 477  
 DB 376 BEYLFVWVDFLVGLVIFATIVGVSMISNMNARAEFOAKIDSIKQYMPKVKTKOLET 435  
 QY 478 RVRTWYETWDSQRMDESLLKTLPTTVQLALAIIDVNFSSISKVDLFKGCOTMIYDML 537  
 DB 436 RVIRWFDLVWANKTKVDEKVLKSLPKLKAIAINVHLDTLKVRIFQDCEAGLLVELV 495  
 QY 538 LRLKSVLVLPDGVCKGEIGKEMYIIKHGEVQVVGPDGTVKLVTLKAGSVFGEISLLA 597  
 DB 496 LKLRPTVSPGDIYCKKGDIGKEMYIINEGKLVV-ADGVTQFVLSGDSYFGEISILN 554  
 QY 598 AGG---GNRRRTANVVAHGFANLLTLDKKTLOEILVHYPDSERILMKKAR-VLLKOK--A 650  
 DB 555 IKGSKSGNRRRTANIRSIGVSLFCLSKDDLMALTEYPEAKKAEKGRQILMKNDLDE 614  
 QY 651 KTAETATPRKDALILFPPEETPKLFTLLGGTGKASLARLL-----KLRQAAQK 702  
 DB 615 ELARAGADPKDL-----BEKVEQLGSSL--DTLQTRFARLLAEYNATQMKMKORLSOLE 666  
 QY 703 KENSEGEE---EGKENEDKOKENEDKOK 728  
 DB 667 SQVKGKDKPLADGEVPGDATK-TEDKQK 694

## RESULT 13

US-09-855-828-14  
 ; Sequence 14, Application US/09855828  
 ; Publication No. US20040137433A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Creech, Christopher D.  
 ; APPLICANT: Jegla, Timothy J.  
 ; APPLICANT: ICAGEN, Inc.  
 ; TITLE OF INVENTION: CNG3B: A Novel Cyclic Nucleotide-Gated Cation Channel  
 ; FILE REFERENCE: 018512-006010US  
 ; CURRENT APPLICATION NUMBER: US/09/855,828  
 ; CURRENT FILING DATE: 2001-05-14  
 ; PRIOR APPLICATION NUMBER: US 60/204,445  
 ; PRIOR FILING DATE: 2000-05-15  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 14  
 ; LENGTH: 694  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: CNGAL  
 US-09-855-828-14

Query Match 20.1%; Score 850; DB 3; Length 694;  
 Best Local Similarity 30.0%; Pred. No. 3e-53;  
 Matches 225; Conservative 149; Mismatches 269; Indels 106; Gaps 22;

QY 27 NEEGSHPNQSOQTAAQENKGEESKLTKTSTPVTSSEPHNTIQDKLSKSSGDI.TTNP 86  
 DB 5 NTQYSHPS-----RTHLKVK-----TSRDLNRAENGLSRAHSSEETS 43  
 QY 87 DPQNAAEPTGVPEOKEMDPGKEGPNSPONKPPAAPVINEYADAOLHNLVKRMORTALY 147  
 DB 44 ---SVLQP-GIAMESTRGLADSGQSFTCQ-----GIARLSRLIFLLRRWAARH 87  
 QY 147 KKLVGDLSSP-----EASQTAKTAVPPVKESDDKPTBHYVRLMLFKVKKMP 197  
 DB 88 VHQDQGPDPDFRFGAELKEVSSQESNAQNVGSEPADRG-----RSAW-----PL 136  
 QY 198 -----TEYLKRIKLPNSI-DSYTDRLVLLMLLVTLAYNNCWFPIRLRVPVOT 246  
 DB 137 AKCNTNTSNTTEEEKTKKDAIIVDPSSNLYRMLTAIALPVFNWYLLICRACFDELQ 196  
 QY 247 ADNIHWLIADIICDIYLYDMLFIQPRLOFVRGGDIIVDSNELRKHYRTSTKFFOLDVAS 306  
 DB 197 SEYLMMLVLDYSADVLYLDVL-VRAATGFLQGLMVSDTNRLWQHVKTTTQFKLDVLS 255  
 QY 307 LIIPDICYLFQFN-PMFRANRMLKTSFFEFNHLESIMDKAYIYRVIRTTGYLLFIH 365  
 DB 256 LVPTDLAYLVKGTNYPEVRFNRLKFSRLEFFDRTETRTNYPNMFRLGNLVLYLIIH 315  
 QY 366 INACVYVWASNYEGIGTRWVYDG-----EGNEYLRCYVAVRTLTITIGLPEPOTL 417  
 DB 316 WNACIYFAISKFIQGTDSWVYPNISIPHGRLSRKIYSLWSTLTTLTTIGETPPVKD 375  
 QY 418 FEIVFOLLNFFSGVVFSSLIQMRDVIQAATANQYFRACDDTIAYNNYSIPKLQVK 477  
 DB 376 BEYLFVWVDFLVGLVIFATIVGVSMISNMNARAEFOAKIDSIKQYMPKVKTKOLET 435  
 QY 478 RVRTWYETWDSQRMDESLLKTLPTTVQLALAIIDVNFSSISKVDLFKGCOTMIYDML 537  
 DB 436 RVIRWFDLVWANKTKVDEKVLKSLPKLKAIAINVHLDTLKVRIFQDCEAGLLVELV 495  
 QY 538 LRLKSVLVLPDGVCKGEIGKEMYIIKHGEVQVVGPDGTVKLVTLKAGSVFGEISLLA 597  
 DB 496 LKLRPTVSPGDIYCKKGDIGKEMYIINEGKLVV-ADGVTQFVLSGDSYFGEISILN 554  
 QY 598 AGG---GNRRRTANVVAHGFANLLTLDKKTLOEILVHYPDSERILMKKAR-VLLKOK--A 650  
 DB 555 IKGSKSGNRRRTANIRSIGVSLFCLSKDDLMALTEYPEAKKAEKGRQILMKNDLDE 614  
 QY 651 KTAETATPRKDALILFPPEETPKLFTLLGGTGKASLARLL-----KLRQAAQK 702  
 DB 615 ELARAGADPKDL-----BEKVEQLGSSL--DTLQTRFARLLAEYNATQMKMKORLSOLE 666  
 QY 703 KENSEGEE---EGKENEDKOKENEDKOK 728  
 DB 667 SQVKGKDKPLADGEVPGDATK-TEDKQK 694

## RESULT 14

US-10-345-680-26  
 ; Sequence 26, Application US/10345680  
 ; Publication No. US20030148394A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Millennium Pharmaceuticals, Inc.  
 ; APPLICANT: Silos-Santiago, Immaculada  
 ; APPLICANT: Venkateswarlu, Karicheti  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
 ; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 1435, 559, 34021, 44099, 25278,  
 ; TITLE OF INVENTION: 641, 260, 55089, 21407, 42032, 46656, 62553, 302, 323,  
 ; TITLE OF INVENTION: 12303, 985, 13237, 13601, 18926, 318, 2038 OR 6351 MOLECULES.  
 ; FILE REFERENCE: MPI02-012P1RM, OMNI  
 ; CURRENT APPLICATION NUMBER: US/10/345,680  
 ; CURRENT FILING DATE: 2003-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/349,511  
 ; PRIOR FILING DATE: 2002-01-18  
 ; PRIOR APPLICATION NUMBER: US 60/360,500  
 ; PRIOR FILING DATE: 2002-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/365,041

```

; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/374,063
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/403,468
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/414,262
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US 60/419,986
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/423,809
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/429,797
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-345-680-26

Query Match      20.1%; Score 850; DB 4; Length 694;
Best Local Similarity 30.0%; Pred. No. 3e-53;
Matches 225; Conservative 149; Mismatches 269; Indels 106; Gaps 22;

QY 27 NEEGSHPSNQSQQTTAQEENKGEESLTKSTPTVTEEPHTNIQDKLSKNSSGDLTTNP 86
DB 5 NTQYSHPS-----RTHLVK-----TSDRLNRAENGLSRAHSSEETS-- 43
QY 87 DPQNAEPTGTVPQEKEMDPGEGNPSQNKPPAAVINEYADAQLHNLVKRMQRTALY 146
DB 44 ---SVLPQ-GIAMETRGADSGGSGFTGQ-----GIARLSRLIFILLRWAARH 87
QY 147 KKKLVGDLSSP-----EASPTAKPTAVPPVKESDDKPTHEYHYLLMFPKVKMPL 197
DB 88 VHQDQGPSPDPFRGAELKEVSSQESNAQNVGQEPADRG-----RSAM-----PL 136
QY 198 -----TEYLKRIKIPNSI-DSYDRLYLLMLLLVTLAYNNWCWFIPLRLVPPYOT 246
DB 137 AKCNTNTSNNTEBEKTKKDAIVDPSSNLYRYRLTAIPALPVFNWYLLICRACFDELQ 196
QY 247 ADNIHWLADIICDIILYDMLFIQPRQFVRGGDIIVDSNELRKHVTSKTFOLDVAS 306
DB \197 SEYLMWLVLVDSADLYLVLDVL-VRARTGFLQGLMVSDTNRLWQHYKTTQFKLDVLVS 255
QY 307 IIPFDICYLFFGFN-PMFRANRLKVTSPFEFNHLESIMDKAYIVRVIRTTGYYLFIH 365
DB 256 LVPTDLAYLVKVTNYEVRNLLKFSRLPEFDRTETRTNYPNMFRIGNLVLYILIIH 315
QY 366 INACVYVWASNYEGIGTTRWYDG-----EGNEYLRCYYWAVRTLITIGGLPEPOTL 417
DB 316 WNAIVFAISKFTGFGTDSWVYPNISIPHGRLSRKYIYSLYVSTLTLTITIGETPPVKD 375
QY 418 FEIVFOLLNFFSGVFFSSLIQMRDVIGNATANYFRACMDDTTAYMNNYSIPKLQVK 477
DB 376 EYLFVVDVFLVGLVLFATIVGVNGSMISNMNASRAEFQAKIDSIKQYWFQFRKVTQDLET 435
QY 478 RVRTWETWDSQRMKLDESLLKTLPTTVQALALADVNFISIISKVDLFGKCDTQMYLDML 537
DB 436 RVIRWFDYLVANKTVDEKEVLSPLDKLAEATINVHLDTLKKVRFQDCEAGLLVELV 495
QY 538 LRLKSLVLYLPDGFVCKGKEIGKEMYYIIKHGEVQVQDGTQVLTAKAGSVFGEISLLA 597
DB 496 LKLRPTVSPGDYICKGDIKEMYYIINEGKLAUV-ADDDGTQFVVLSDGSYFGEISILN 554
QY 598 AGG---GNRTANVAVGAFANLLTLDKKTQLEILVHPDPSERILMKAR-VLLKQK---A 650
DB 555 IKGSKGNRRNTAIRSIGYDGLFCLSKDLDMEALTEYPEAKKALAEKGRQILMKDLIDE 614
QY 651 KTEATPPRKDALLPPPKETPKLFTKLLGGTGKASLRL-----KLKREAAQK 702
DB 615 ELARAGADPKDL-----EBKVFOLGSSL---DTLQTRFARLLAEYVATQMKQRLSQLE 666

```

```

QY 703 KENSEGEE---EGKENEDKOKENEDKOK 728
DB 667 SQVGGGDKPLADGEVPCDATK-TEDKQO 694

RESULT 15
US-10-174-333-75
; Sequence 75, Application US/10174333
; Publication No. US20040029220A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Zethusen, Bryan D.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Smithson, Glennda
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Grose, William M.
; APPLICANT: Szekeres, Edward S.
; APPLICANT: Alsobrook, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Li, Li
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-783 CIP1
; CURRENT APPLICATION NUMBER: US/10174,333
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 60/193,664
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/194,614
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,063
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,066
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,067
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,068
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,069
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,070
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,510
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/219,855
; PRIOR FILING DATE: 2000-07-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: Curaseqdist version 0.1
; SEQ ID NO 75
; LENGTH: 694
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-174-333-75

Query Match      20.1%; Score 850; DB 4; Length 694;
Best Local Similarity 30.0%; Pred. No. 3e-53;
Matches 225; Conservative 149; Mismatches 269; Indels 106; Gaps 22;

QY 27 NEEGSHPSNQSQQTTAQEENKGEESLTKSTPTVTEEPHTNIQDKLSKNSSGDLTTNP 86

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2005, 19:54:15 ; Search time 13 Seconds  
(without alignments)  
443.837 Million cell updates/sec

Title: US-09-855-828-1  
Perfect score: 4234  
Sequence: 1 MFPSLTKVKNVKPIGENNEN.....PSAEGGEVLTIETVKEKAKQ 809

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 53982 seqs, 7132107 residues

Total number of hits satisfying chosen parameters: 53982

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New.\*  
1: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/US11\_NEW\_PUB pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB pep.\*

\*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	818.5	19.3	690	6	US-10-613-744-19
2	484	11.4	261	6	US-10-613-744-18
3	317.5	7.5	1159	6	US-10-613-744-12
4	302.5	7.1	857	6	US-10-613-744-11
5	282.5	6.7	597	6	US-10-613-744-7
6	131.5	3.1	3375	7	US-11-044-111-23
7	124	2.9	447	6	US-10-793-626-2900
8	124	2.9	450	6	US-10-793-626-3226
9	117	2.8	57	6	US-10-613-744-38
10	116.5	2.8	410	6	US-10-821-234-912
11	113	2.7	4374	7	US-11-128-572-2
12	112	2.6	635	6	US-10-821-234-1573
13	109.5	2.6	2897	6	US-10-499-715-2
14	107	2.5	57	6	US-10-613-744-39
15	106.5	2.5	598	7	US-11-113-837-18
16	106.5	2.5	644	6	US-10-793-626-1436
17	106.5	2.5	665	7	US-11-113-837-19
18	105.5	2.5	440	7	US-11-108-172-1059
19	105	2.5	1618	6	US-10-984-645-2
20	104	2.5	567	6	US-10-485-517-216
21	104	2.5	635	7	US-11-113-837-16
22	104	2.5	877	6	US-10-485-517-200
23	103.5	2.4	482	6	US-10-821-234-1413
24	102.5	2.4	547	6	US-10-770-726-87
25	102.5	2.4	667	6	US-10-821-234-1477

26	102	2.4	1390	7	US-11-063-343-35	Sequence 35, Appl
27	101.5	2.4	4868	7	US-11-044-111-24	Sequence 24, Appl
28	101	2.4	769	6	US-10-485-517-401	Sequence 401, Appl
29	101	2.4	1116	6	US-10-485-517-238	Sequence 238, Appl
30	101	2.4	1117	6	US-10-485-517-206	Sequence 206, Appl
31	100.5	2.4	795	6	US-10-770-726-49	Sequence 49, Appl
32	100.5	2.4	1122	6	US-10-821-234-1657	Sequence 1657, Ap
33	100	2.4	559	6	US-10-793-626-1376	Sequence 1376, Ap
34	99.5	2.4	457	6	US-10-982-545-8	Sequence 8, Appl
35	99.5	2.4	457	6	US-10-982-545-13	Sequence 13, Appl
36	99.5	2.4	715	7	US-11-089-551A-47	Sequence 47, Appl
37	98.5	2.3	290	6	US-10-821-234-862	Sequence 862, Appl
38	98.5	2.3	550	7	US-11-113-837-17	Sequence 17, Appl
39	98.5	2.3	853	6	US-10-821-234-1110	Sequence 1110, Ap
40	97.5	2.3	499	6	US-10-508-263-94	Sequence 94, Appl
41	97.5	2.3	756	7	US-11-188-743-20	Sequence 20, Appl
42	97.5	2.3	828	6	US-10-501-039-2	Sequence 2, Appl
43	97	2.3	313	6	US-10-873-528-163	Sequence 163, Appl
44	97	2.3	1263	6	US-10-485-517-127	Sequence 127, Appl
45	97	2.3	2261	6	US-10-995-561-600	Sequence 600, Appl

ALIGNMENTS

RESULT 1

US-10-613-744-19  
; Sequence 19, Application US/10613744  
; Publication No. US2005027093A1  
; GENERAL INFORMATION:  
; APPLICANT: MacKinnon, Roderick  
; APPLICANT: The Rockefeller University  
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With  
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation  
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof  
; FILE REFERENCE: 018512-002901US  
; CURRENT APPLICATION NUMBER: US/10/613, 744  
; CURRENT FILING DATE: 2003-07-03  
; PRIOR APPLICATION NUMBER: US/09/275,252  
; PRIOR FILING DATE: 1999-03-24  
; PRIOR APPLICATION NUMBER: US 09/045,529  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: US 09/054,347  
; PRIOR FILING DATE: 1998-04-02  
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307  
; PRIOR FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 690  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-613-744-19

Query Match	19.1%	Score	818.5	DB	6	Length	690
Best Local Similarity	30.0%	Pred. No.	1.2e-51				
Matches	221	Conservative	128	Mismatches	262	Indels	125
Gaps	22						
Qy	5	LTKVNVKPIGENNENEQSSRRNEEGSHPS	---	NOSQOTTAOENKGE	-----	KSL	53
Db	17	VTWPNVIVP	-----	DIEKEIRRMENACSFSDDDSDAYS	TESEENENPHARGCSFYKSL	71	
Qy	54	KTGSTPTVSEPHNTIQDKLKNKNSGDLTTNPDPONAAEPTGTVP	EQKEMDPCKEGPN	113			
Db	72	R-KGQPSQREQLPGAIAIFNVNNS	-----	NKD	---	QEPPEKKKKKKKSKSDDKNE	121
Qy	114	PQNKPAAPVINEYADAQLHNLVKRMQRQTALVKKLVGCDLSSPEAS	POTAPVPPV	173			
Db	122	NKNDPE	-----	KKKKKKKKKKK	-----	E	141
Qy	174	KESDDKPTHEYRLWFKVKMKMPLTEYLKRIKLPNSIDSYDRLYL	LWLLVTL	231			
Db	142	EKSKDKKEH	-----	KKEVVV	-----	IDP-SGNTYYNWLFCITL	180



Db 484 SHPGRIAVHYFKGW-FLIDMVAAPDL--LIFGSG-----SEELIGL 523  
Qy 345 MDKAYIYRVRTGYL-----LFIL-----HINACVYWAGNYE-----CI 380  
Db 524 LKTBALRLRVRVARKLDYSEYGAVALFLMCTFALIAHWLACIWAIGNMOPHMSRI 583  
Qy 381 GTTRWV-----YDGE-----NEYLRCCYVAVRTLITIG-GLPEPOTLFEIV 422  
Db 584 G---WLHNLGDOQKGPYNSGLGPGSIKOKYVVALYFTFSSLTSGVGNVSPNTSEKIF 640  
Qy 423 QLLNPFSGVVFSSLLIGOMRDVIGATANQNYFRACMDDTIAYMNNYSIPKLVOKEVRTW 482  
Db 641 SICWMLIGSLMYASIFGNYSIAIRQYSGTARYHTQMLRVREIFRHOJPNLQRLSEY 700  
Qy 483 YEYTWDSQRLMDESLLKTLPTTVQALALADVNFSIIISKVDLFGKCDTQMIYDMLRLKS 542  
Db 701 FQAWSYTNGIDMNAVLKGFPECLQADICHLNRSLLQHCXPRGATKGLRALANKPKT 760  
Qy 543 VLYLPGDFVCKGKEIGKEMYIIKHGEVQVVGPGDGTKVLVTLKAGSVFGE-ISLLAAGGG 601  
Db 761 THAPPGDTLVHAGDLLTALYFISRGSIEIL---RGDWVVALGKNDIFGEPLNLYARPG- 816  
Qy 602 NRTANVAVHGFANLTLTKTLQELVHYPD 633  
Db 817 -KSGDVRALTYCDLHKIHRDDLLEVLDMYPE 847

## RESULT 4

US-10-613-744-11  
; Sequence 11, Application US/10613744  
; Publication No. US20050272093A1  
; GENERAL INFORMATION:  
; APPLICANT: MacKinnon, Roderick  
; APPLICANT: The Rockefeller University  
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With  
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation  
; TITLE OF INVENTION: Cation Channel Proteins, and Uses Thereof  
; FILE REFERENCE: 018512-002901US  
; CURRENT APPLICATION NUMBER: US/10/613,744  
; CURRENT FILING DATE: 2003-07-03  
; PRIOR APPLICATION NUMBER: US/09/275,252  
; PRIOR FILING DATE: 1999-03-24  
; PRIOR APPLICATION NUMBER: US 09/045,529  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: US 09/054,347  
; PRIOR FILING DATE: 1998-04-02  
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 857  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-613-744-11

Query Match 7.1%; Score 302.5; DB 6; Length 857;  
Best Local Similarity 23.2%; Pred. No. 2.2e-14;  
Matches 110; Conservative 91; Mismatches 226; Indels 47; Gaps 15;  
Qy 202 KRKLPSIDSYDRLYLLMLLLVTLAYNWCNCFIPLRLVF-----PYQTADNIHWLI 255  
Db 41 RVKLRFVVSYPDHKYRIEAFVLVAVYTAWSPEFGFLRPRPLSLTDNI---VN 97  
Qy 256 ADICDIYLYDMLFIQRIQFVRGGDIIYVDSNELKHYRTSTKFDQDVASIIIPD---- 311  
Db 98 AFFAIDIIMTFVGYLDKSTYL-----IVDRKQIAFKYLRSW-FLDLVSTIPSEAMR 151  
Qy 312 ICYLPFGFNPMFRANMLKYTSFPFNNHLESIMDKAYIY-RVIRTGTYLLFILHINACV 370  
Db 152 ISSOSYGLFNMLRLWRLRRVGALFA---RLEKDRNFYFWRCAKLVCVTLFAVHCAACF 208

Qy 371 YW-----ASNVEGIGTTRWVYDGEENYLRCYYWAVRTLITIG-CLPEPOTLFEIV 421  
Db 209 YYLIAARNSNPACTWIGANVANFLEESLWMRYVVTSMYSITTLTVGYDGLHPVNTKEMI 268  
Qy 422 FOLLNPFSGVVFSSLLIGOMRDVIGATANQNYFRACMDDTIAYMNNYS-----PKLVOK 477  
Db 269 FDFYFMLNLGLTAYLIGNMTNLVVHGTSTRNFR-----DTIOAASNFARHNLPPRLQD 324  
Qy 478 RV--RTWVEYTWDSQRLMDESLLKTLPTTVQALALADVNFSIIISKVDLFGKCDTQMIYD 535  
Db 325 QMLAHLCLKYRTDSEG-LQQOETLDALPKAIRSSISHFLFYSLMDKVLYLFRGVNSNDLLFQ 383  
Qy 536 MLRLKSVLYLPGDFVCKGKEIGKEMYIIKHGEVQVVGPGDGTKVLV-TLKAGSVFGEIS 594  
Db 384 LVSEMKAEYFPPEKDVILQNEAPTDFYILVNGTADLVVDVTGTESTIVREVKAGDIIIGEIG 443  
Qy 595 LLAAGGGRNRTANVAVHGFANLTLTKTLQELVHYPDSEIRILMKKARVLLKO 648  
Db 444 VLICY---RPQLFTVTRTKRLCQLLRNRTTFLNIQANVGDGTIIMNNLLOHLKE 494

## RESULT 5

US-10-613-744-7  
; Sequence 7, Application US/10613744  
; Publication No. US20050272093A1  
; GENERAL INFORMATION:  
; APPLICANT: MacKinnon, Roderick  
; APPLICANT: The Rockefeller University  
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With  
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation  
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof  
; FILE REFERENCE: 018512-002901US  
; CURRENT APPLICATION NUMBER: US/10/613,744  
; CURRENT FILING DATE: 2003-07-03  
; PRIOR APPLICATION NUMBER: US/09/275,252  
; PRIOR FILING DATE: 1999-03-24  
; PRIOR APPLICATION NUMBER: US 09/045,529  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: US 09/054,347  
; PRIOR FILING DATE: 1998-04-02  
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 597  
; TYPE: PRT  
; ORGANISM: Paramesicium tetraurelia  
US-10-613-744-7

Query Match 6.7%; Score 282.5; DB 6; Length 597;  
Best Local Similarity 23.3%; Pred. No. 3.8e-13;  
Matches 118; Conservative 95; Mismatches 212; Indels 81; Gaps 23;  
Qy 178 DKPT-----BHYI-RLNLFVKVKNPLTEYL-----KRKLPSIDSYDRLYLLMLLLVTL 227  
Db 98 EKPTGFGQGHFLNQLTLPLRKLKLYKCGNQKLMRPELSASIMWNIYIITIL----- 153  
Qy 228 AYNNWCNCFIPLRLVFPVQATADNIHWLIADIICDIYLYDMLFIQRIQF-----VRGCDI 283  
Db 154 --NINVLVSIKIAFKAFDEOSQDDFVQARQIIFDVLPSSVFM-LEILLKENTCYCYKCAV 210  
Qy 284 IVDSNELKHYRTSTKFDQDVASIIIP-----FDICYLFFGFPNPMFRANMLKYTSFPF 337  
Db 211 IENRYQIAKNYLRS--FFDFIVVIFVIFSLRFDQLYDLVI-----ILKVFQITKF 261  
Qy 338 NHLESIMDKAYIYRVI---RTTGY-LLFILHINACVYWASNYEGIGTTRWV-----YDG 389  
Db 262 SRNLFORLELTAIQIVIDLVLKGYTILAAAHFSACTIWFVLVSGTNPNDTSWIKAOINEN 321  
Qy 390 EG--NEVLRCCYVAVRTLITIG-GLPEPOTLFEIVFOLLNPFSGVVFSSLLIGOMRDVIG 446  
Db 322 EQWFNOYLSHLSYIITMTTIGYDITPQNLRRERVFAVGMALSAVGVFSGYSGINSIYA 381

```
QY 447 AATANQNYFRACDDTIAYMNNYSIPKLVQKRVRTWYETWDSQRMLESLLK---TLP 503
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
382 EMSRQSFQIRTDNNLKFKIRIKINGINKHAEKIRKYFEYVWSDQMEDNDREVYKSEMP 441
QY 504 TTVOLALADVNFSIISK-----VDLPKGCOTQMIYDMLRLKSVL-----YLPGDF----- 550
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
442 KOLAEMKIDTNNKIQKNSFLVNNFS-----EQFLISLSKVLIIEKYVPESTIYLVK 494
QY 551 ---VCKKGEIGKEMVIKIHGEVQ---VLGGPDGT-KVLVTLK-AGSVFGEISLLAAGGN 602
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
495 LINILQNDPSNYLXLSNGLSFYITLNNKQOTIKVLETIKNEGAQFVLEFFQS---Q 551
QY 603 RRTANVAHGFAANLTLDKKTIQEIL 628
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
552 AYQVSCSKNQFSVVLKIDKSKQPMWII 577

RESULT 6
US-11-044-111-23
; Sequence 23, Application US/11044111
; Publication No. US20050272362A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Wen
; APPLICANT: Strasbourg, Gale
; APPLICANT: Linz, John
; TITLE OF INVENTION: Genetic Test for PSP-Susceptible Turkeys
; FILE REFERENCE: MSU-09308
; CURRENT APPLICATION NUMBER: US/11/044,111
; CURRENT FILING DATE: 2005-01-27
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 3375
; TYPE: PRT
; ORGANISM: Meleagris gallopavo
US-11-044-111-23

Query Match 3.1%; Score 131.5; DB 7; Length 3375;
Best Local Similarity 21.3%; Pred. No. 0.29;
Matches 109; Conservative 66; Mismatches 198; Indels 139; Gaps 22;

QY 322 MFRANMLKYTSFFFNHLESIMDKAYIRVIRTYGILLFLHINACVYVWASVYEGIG 381
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2485 LFRLLQLLCEGHNDNFQNYLRTQTGNTTTINIICTVDYLLRLQESISDFYV--YVSG-- 2540
QY 382 TTRWYVDGEGNEVLCRYVAVRITLITIGLPEPQTLFEIVFQLLNFFSGVFVFSILGQM 441
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2541 --KVIDEGOKR-----NFSKAMAVAKQVFNSL 2566
QY 442 RDVI-GAATANQNYF--RACMDDTIAYMNNYS--IPKLVQKRVRTWYETWDSQRMLES 496
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2567 TEYIQGPCTGNQQSLAHSRLWDVAVGFLVFAHMMKLAQ-----DSSQIGLLK 2615
QY 497 DLLKTLPTTVQALAI-----DVNFSIISKV---DLFKGCDTOMI---YDMLRLKSVL-- 544
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2616 ELDDIQDMVMVLLSLLEGNVVNGTIARQVDMPVNESSNVTMLKFFDMFLKLRDIVAS 2675
QY 545 -----YL--PGDFVCKKGIGKEMVIIKHGEVQVVLGGPDGTKVLTALKAGSVFGISLLA 597
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2676 DAFRDIYVDPRLGISKK-DFOKAMDSQKQYE-----PSEVQFLIISCSSEADNEMIDVEA 2728
QY 598 AGGNGRNTANVAHGFAANLTLDDKTKLQELVHYPDSEIRLMKKARVLLKQAKTAEATP 657
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2729 FVGRSQEPARDIGFNVAVLLT-----NLAHVHVDQRL-----RTFLEQAASILE--- 2773
QY 658 FRKDLALLFPKPEETPKFLTKLLGGTGKASLARLL-KLKREQAOKKENSEGEGEBEGKE- 715
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2774 -----YFRPFLGRIEINGAARRIERLYFEISAANKAQWEMPQDGGKEV 2816
QY 716 NEDKOKENEDKOKENEDKOKREPEEKPLDRPE--CTASPIAVEEPEHPSVRT 773
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2817 VEPPQKEEAAAETEKADEKADTENGEGKDGGAEGGPEVETPEKQOKASPPRERKEP----- 2870
```

```
QY 774 VLPRGTSRQSLIISMAPSAEGGEEVLT-IEVK 804
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
2871 -----PPPEGAFETWTELEVO 2886

RESULT 7
US-10-793-626-2900
; Sequence 2900, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2900
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-2900

Query Match 2.9%; Score 124; DB 6; Length 447;
Best Local Similarity 27.9%; Pred. No. 0.067;
Matches 48; Conservative 24; Mismatches 72; Indels 28; Gaps 6;

QY 662 LALLFPKPEETPKFLTKLLGGTGKASLARLLKLKREQAOKKE----- 704
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MANLDERKKEITANLISKAQEAVERGDLLETARNLKADIDAQKREVEELEQLSKEIEASAPK 60
QY 705 NSEGEEREGKEDKQ-----KENEKOKENEDKGENEDK-DKGREPEEKPLDRPECTA 758
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 QDEPPKDEGAEDNKDNGSGESENKPSDDPEGTSDEEKPDPAKPDCKPEETPE-Tp 119
QY 759 SPIAVEE--EPHSVRRTVLPRTGTSRQSLIISMAPSAEGGEEVLTIEVKEKAK 808
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
120 TIEKVEEPTBELKKEKDKKEGAKRSMAKLQNP--ETNEBILAFEQYMKSK 169

RESULT 8
US-10-793-626-3226
; Sequence 3226, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3226
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-3226

Query Match 2.9%; Score 124; DB 6; Length 450;
Best Local Similarity 27.9%; Pred. No. 0.067;
Matches 48; Conservative 24; Mismatches 72; Indels 28; Gaps 6;
```

```

; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-912

Query Match          2.8%; Score 116.5; DB 6; Length 410;
Best Local Similarity 24.0%; Pred. No. 0.21;
Matches 55; Conservative 29; Mismatches 102; Indels 43; Gaps 7;

Qy 426 NFESGVFVSSLLIGQMRDVGAAATANONYFRACMDDTIAYMNN----- - 468
Db 177 DIFDAMFSVSFIAGE--TWIQOQDEGDNFVVIDOGETDVVYVNNWATSVGEGGSGFELAL 234
Qy 469 -YSIPKLVQKRVRT-----WYEYTWDSQRMLEDSDLLKLTPTTVQLALADVNFSIISKVD 523
Db 235 IYGTTPRAATVKAKTNVKLMGIDRDSYRRLMGSTLRK-----KMYEELFSKVS 283
Qy 524 LFXGCDTQMIYDMLLRKLSVLYLPDGFVCKKGIGKEMVIIKHGEVQVLGGPDGDKVLV- 582
Db 284 ILESLOKWERLTVADALEPVOFEDGQKIVVQEGPGEFFILLEGSAVLORRSENEEFVE 343
Qy 583 --TLKAGSVGEISLLAAGGNN-RTANVVAHGFAANLLTLDDKKTLOEIL 628
Db 344 VGRLGPSDYGEITALLM---NRPRATVVARGPLKCVKLDRRPRFVL 388

RESULT 11
US-11-128-572-2
; Sequence 2, Application US/11128572
; Publication No. US20050266513A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim International GmbH
; TITLE OF INVENTION: Method for identifying compounds that have the potential
; TITLE OF INVENTION: activity of Myc
; FILE REFERENCE: Case 12-0231-PCT
; CURRENT APPLICATION NUMBER: US/11/128,572
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 4374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-128-572-2

Query Match          2.7%; Score 113; DB 7; Length 4374;
Best Local Similarity 25.6%; Pred. No. 9;
Matches 33; Conservative 25; Mismatches 41; Indels 30; Gaps 4;

Qy 635 ERILMKKARVLLKQAKTAPATPRKDALLFPKKEETPKFLTLLGGTGKASLARLLKL 694
Db 2379 EDVLMDEAPNSLSQ--ASTLQAN--REDSMNILOPEDE----- 2412
Qy 695 KREQAAQKKENSEGEEGKENEKDKENEDKQKENEKDKGKREPEEKPLDR- 753
Db 2413 --EHTQEDSSGSGNEDEDDSDQEEEEEEEEDEDDQDEDEGEGEDDDDDGSEMEDJED 2470
Qy 754 -PECTASPI 761
Db 2471 YPDNNVASPL 2479

RESULT 12
US-10-821-234-1573
; Sequence 1573, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821a;

```



Query Match	2.5%;	Score 106.5;	DB 7;	Length 598;
Best Local Similarity	17.9%;	Pred. No. 1.8;		
Matches 102;	Conservative 102;	Mismatches 212;	Indels 155;	Gaps 27;
Qy	16	ENNENEQSSRRRNEGSHSPNSOQTTAQEENKGEBSLTKTSTPTVTSBEPTNIOQKLSK	75	
Db	60	QONHNDPMTGLAHQFAHQNIAGRASPYSGRG-----PSPAQRPRTS---GNSGQ	107	
Qy	76	KNSGSDLTTNDPQNAAEPTGTVPQKEMDPQKEGPNSPQNKPAAAPVINEYADAQLHNL	135	
Db	108	QQTGYNLSALPMPSNTQTEFAPAPER---NPDKYGNANNNQKCSQLASDF---FKDS	160	
Qy	136	VKRWNR---ORTALYKKKLVE-GDLSPPASQTA--KPTAVPVVKESDDKPTHEYHYRL--	187	
Db	161	VKRARENRQORQSEMQKLGETNDARRRRSIWSTAGRKEGQYLRFURTKDKP-ENYQTIKI	219	
Qy	188	-----LWPKVKOMPTLEYLKRILKPLNSIDSYTDRLYLLWLLVLTAY	229	
Db	220	IGKGAFGEVKLVQKADGKVKAMKSLIKTEMPKQOLAH-VRAERD-----ILAE	268	
Qy	230	NWNCWFIPLRLVFPYQTADNIHYWLIADIICDILYLDMLFIQRLQFVRGGDII-----	284	
Db	269	SDSPWVKLYTTF-----QDANFLYMLMEFLPGCDLMTMLIK	305	
Qy	285	--VDNSELRKHRYRTSTKQFDVASIIPDICYLFFGF-----NPMFRANRMLUKYTSF	334	
Db	306	YEIPSEDIRTFYASIVLAIDAIVHKI-----GFTHRDIKPONILLDRGHGVKLTFD	356	

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2005, 15:46:45 ; Search time 1330 Seconds  
(without alignments)  
12176.844 Million cell updates/sec

Title: US-09-855-828-3  
Perfect score: 2430  
Sequence: 1 atgtttaatcgtgacaaa.....aagaaaagctgaagcaataa 2430

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_21.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2430	100.0	2757	6	AAD25729 Human cyc
2	654.4	26.9	5	AAS69653	Aas69653 DNA encod
3	579	23.8	4380	13	ACN42631 Human dia
4	577.4	23.8	2607	8	ABZ80565 Nucleotid
5	573.6	23.6	3811	5	AAS91658 DNA encod
6	569.6	23.4	4238	10	ADB33355 Primary r
7	568	23.4	3236	13	ADS64556 Norway ra
8	518	21.3	2244	5	AAS76156 DNA encod
9	518	21.3	2244	5	AAS76184 DNA encod
10	223	9.2	3275	4	ABL19457 Drosophil
11	176.4	7.3	2500	10	ADD93243 Cyclic nu
12	176.4	7.3	2500	13	ADR4566 Breast ca
13	160.6	6.6	289	5	AAS75291 DNA encod
14	136.2	5.6	2109	4	ABL18085 Drosophil
15	133.2	5.5	3027	10	ADF56505 Modified
16	132.8	5.5	2615	12	ADQ67343 Novel hum
17	132.8	5.5	3486	10	AAD58492 Human CNG
18	131.6	5.4	1995	8	AAD56241 Rat CNG c
19	131.6	5.4	1995	8	AAD56239 Rat CNG c

20	131.6	5.4	1995	14	ADZ88276	Adz88276 Rat CNGA2
21	131.6	5.4	1995	14	ADZ88280	Adz88280 Mutated R
22	131.6	5.4	3027	10	ADF56506	Adf56506 Modified
23	131.6	5.4	3027	10	ADF56508	Adf56508 Rat olfac
24	129	5.3	526	4	ABA61898	Abas61898 Human foe
25	129	5.3	526	4	AAI41823	AAi41823 Probe #10
26	129	5.3	526	4	AAK36108	Aak36108 Human bon
27	129	5.3	526	4	AAK10209	Aak10209 Human bra
28	129	5.3	526	4	ABS35801	Abss35801 Human liv
29	128.4	5.3	1995	8	AAD56242	Aad56242 Rat CNG c
30	128.4	5.3	1995	8	AAD56240	Aad56240 Rat CNG c
31	128.4	5.3	1995	14	ADZ88282	Adz88282 Mutated R
32	128.4	5.3	1995	14	ADZ88278	Adz88278 Mutated R
33	128	5.3	2877	10	ADF56507	Adf56507 Modified
34	127	5.2	127	4	ABA74393	Abas74393 Human foe
35	127	5.2	127	4	AAI54858	AAi54858 Probe #23
36	127	5.2	127	4	AAK49030	Aak49030 Human bon
37	127	5.2	127	4	AAK22857	Aak22857 Human bra
38	127	5.2	127	4	ABS48687	Abss48687 Human liv
39	124.4	5.1	5133	4	ABU14931	Abu14931 Drosophil
40	120.6	5.0	5692	4	ABU19456	Abu19456 Drosophil
41	116.8	4.8	1989	8	ABX94804	Abx94804 Bovine CN
42	116.8	4.8	3166	13	ADS64555	AdS64555 Bovine cy
43	115.2	4.7	1989	8	ABX94805	Abx94805 Bovine CN
44	115.2	4.7	1992	13	ADS64558	AdS64558 Cyclic nu
45	113.6	4.7	1989	8	ABX94806	Abx94806 Bovine CN

ALIGNMENTS

RESULT 1

AAD25729  
ID AAD25729 standard; DNA; 2757 BP.

XX AC AAD25729;

XX DT 26-MAR-2002 (first entry)

XX DE Human cyclic nucleotide-gated cation channel 3 beta subunit DNA.

XX KW Human; cyclic nucleotide-gated cation channel 3 beta subunit; CNG3B;  
cyclic nucleotide-gated ion channel; contraceptive; vision disorder;  
male infertility; genetic defect; reporter-ligand interaction; CNG;  
viral infection; cancer; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT misc\_feature 1..49

FT FT /note= a

FT FT /note= "These bases are absent in the sequence shown as

FT CDS SEQ ID NO:2 in Page 64-65 of the specification"

FT FT 112..2541

FT FT /tag= b

FT FT /product= "Human CNG3B protein"

FT FT /note= "This region is specifically claimed in claim 4 as

FT misc\_feature SEQ ID NO:4"

FT FT 2753..2757

FT FT /tag= c

FT FT /note= "These bases are absent in the sequence shown as

FT FT SEQ ID NO:2 in Page 64-65 of the specification"

XX PN WO200188090-A2.

XX PD 22-NOV-2001.

XX XX 15-MAY-2001; 2001WO-US015814.

XX XX 15-MAY-2000; 2000US-0204445P.

XX PR 14-MAY-2001; 2001US-00855828.

XX XX (ICAG-) ICAGEN INC.

[illegible]

1732	AAATCCGGTCTCTATTTGGCTCGTGACTTTGTGCTGCAAAAGGGAGAAATTGGCAAGGAA	1791
	DB	
1681	ATGTATATCATCAAGCATGGAGAAGTCCAAGTTCTTTGGAGGCCCTCATGTGACTAAAGTT	1740
	QY	
1792	ATGTATATCATCAAGCATGGAGAAGTCCAAGTTCTTTGGAGGCCCTCATGTGACTAAAGTT	1851
	DB	
1741	CTGGTTACTCTGAAAGCTGGGTCCGGTGTTTGGAGAAATCAGCCTTCTAGCAGCAGGAGGA	1800
	QY	
1852	CTGGTTACTCTGAAAGCTGGGTCCGGTGTTTGGAGAAATCAGCCTTCTAGCAGCAGGAGGA	1911
	DB	
1801	GGAAACCGTTCGAATCGTCAATGTGGTGGCCACGGGTTTGCCTAATCTTTTAACTCTAGAC	1860
	QY	
1912	GGAAACCGTTCGAATGTGGTGGCCACGGGTTTGCCTAATCTTTTAACTCTAGAC	1971
	DB	
1861	AAAAGACCCTCCAAAGAAATTTCTAGTGCATTATCCAGATTCTGAAAGGATCCCTCATGAAG	1920
	QY	
1972	AAAAGACCCTCCAAAGAAATTTCTAGTGCATTATCCAGATTCTGAAAGGATCCCTCATGAAG	2031
	DB	
1921	AAAGCCAGAGTGCTTTTAAAGCAGAGGCTTAAGACCGCAGAGAACAACCCCTCCAAAGAAA	1980
	QY	
2032	AAAGCCAGAGTGCTTTTAAAGCAGAGGCTTAAGACCGCAGAGAACAACCCCTCCAAAGAAA	2091
	DB	
1981	GATCTTTGCCCTCTCTTTCCACCGAAAGAGAGACACCCAAAATGTTTTAAAACTCTCCTA	2040
	QY	
2092	GATCTTTGCCCTCTCTTTCCACCGAAAGAGAGACACCCAAAATGTTTTAAAACTCTCCTA	2151
	DB	
2041	GGAGGCA CAGGAAAAGCAAGTCTTTGCAAGACTACTCAAATTTGAAGCGGAGCAGCAGCT	2100
	QY	
2152	GGAGGCA CAGGAAAAGCAAGTCTTTGCAAGACTACTCAAATTTGAAGCGGAGCAGCAGCT	2211
	DB	
2101	CAGAGAAAAGAAAATTTCTGAAGGAGGAGAGAGAGAGAAAATGAAGATTAACAA	2160
	QY	
2212	CAGAGAAAAGAAAATTTCTGAAGGAGGAGAGAGAGAGAAAATGAAGATTAACAA	2271
	DB	
2161	AAAGAAAATGAAGTAAACAAAAGAAAATTCAGATAAAGGAAAGAAAATGAAGATAAA	2220
	QY	
2272	AAAGAAAATGAAGTAAACAAAAGAAAATTCAGATAAAGGAAAGAAAATGAAGATAAA	2331
	DB	
2221	GATAAAGGAAGAGAGCCAGAGAAGAGAGCCACTGGACAGACTCTGAATCTCAGCAAGTCCT	2280
	QY	
2332	GATAAAGGAAGAGAGCCAGAGAAGAGAGAGCCACTGGACAGACTCTGAATCTCAGCAAGTCCT	2391
	DB	
2281	ATTGCAGTGGAGGAAGAACCCCACTCAGTTTAGAGGACAGTTTTTACCAGAGGGACTTCT	2340
	QY	
2392	ATTGCAGTGGAGGAAGAACCCCACTCAGTTTAGAGGACAGTTTTTACCAGAGGGACTTCT	2451
	DB	
2341	CGTCAATCACTCATTTATCAGCATGGCTCCTTTCTGCTGAGGCGGAGAGAGGTTTCTTACT	2400
	QY	
2452	CGTCAATCACTCATTTATCAGCATGGCTCCTTTCTGCTGAGGCGGAGAGAGGTTTCTTACT	2511
	DB	
2401	ATTGAAGTCAAGAAAAGGCTTAAGCAATAA	2430
	QY	
2512	ATTGAAGTCAAGAAAAGGCTTAAGCAATAA	2541
	DB	

## RESULT 2

RECORD 2  
AAS69653  
ID AAS69653 standard: cDNA: 2516 BP.

AC AAS69653;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #5457.

Human; chromosome mapping; gene mapping; gene therapy; forensic; KW  
food supplement; medical imaging; diagnostic; genetic disorder; ss. KW

OS Homo sapiens.

PN WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US008631.  
31-MAR-2000; 2000US-00540217.  
23-AUG-2000; 2000US-00649167.  
(HYSE-) HYSEQ INC.  
Drmanac RT, Liu C, Tang YT;  
WPI; 2001-639362/73.  
P-PSDB; ABG05466.  
New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity.  
Claim 1; SEQ ID NO 5457; 103pp; English.  
The invention relates to isolated polynucleotide (I) and polypeptide (II)  
sequences. (I) is useful as hybridisation probes, polymerase chain  
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
and in recombinant production of (II). The polynucleotides are also used  
in diagnostics as expressed sequence tags for identifying expressed  
genes. (I) is useful in gene therapy techniques to restore normal  
activity of (II) or to treat disease states involving (II). (II) is  
useful for generating antibodies against it, detecting or quantitating a  
polypeptide in tissue, as molecular weight markers and as a food  
supplement. (II) and its binding partners are useful in medical imaging  
of sites expressing (II). (I) and (II) are useful for treating disorders  
involving aberrant protein expression or biological activity. The  
polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
coding sequences of the invention. Note: The sequence data for this  
patent did not appear in the printed specification, but was obtained in  
electronic format directly from WIPO at  
ftp.wipo.int/pub/published pct sequences

Query Match 26.9%; Score 654.4; DB 5; Length 2516;  
Best Local Similarity 99.1%; Pred. No. 1.5e-133;  
Matches 658: Conservative 0; Mismatches 6; Indels 0; Caps 0;

1767	Qy	GTTTGGAGAAATCAGCCTTCTACAGCAGAGAGAGAAACCGTCGAACCTGCCAATGTGGT	1826
1853	Db	GTCAAGAAAACAACAGCCTTCTAGCAGCAGAGAGAGAAACCGTCGAACCTGCCAATGTGGT	1912
1827	Qy	GGCCCAAGGTTTGGCCAACTCTTTAACTCTAGACAAAAAGACCCTCCAGAAAAATCTTAGT	1886
1913	Db	GGCCCAAGGTTTGGCCAACTCTTTAACTCTAGACAAAAAGACCCTCCAGAAAAATCTTAGT	1972
1887	Qy	GCATTATCCAGATTCTGAAAGGATCTCTGTAAGAAAGCCAGAGTGCCTTTTAAAGCAGAA	1946
1973	Db	GCATTATCCAGATTCTGAAAGGATCTCTCATGAAGAAAGCCAGAGTGCCTTTTAAAGCAGAA	2032
1947	Qy	GGCTAAGACCGCAGAGCAACCCCTCCAGAAAAAGATCTTTGCCCTCTCTTCCCAACCGAA	2006
2033	Db	GGCTAAGACCGCAGAGCAACCCCTCCAGAAAAAGATCTTTGCCCTCTCTTCCCAACCGAA	2092
2007	Qy	AGAAGAGACACCCNAACTGTTTAAAACTCTCTTAGAGGGCACAGGAAAAAGCAAGTCTTGC	2066
2093	Db	AGAAGAGACACCCAAAATGTGTTTAAAACTCTCTTAGAGGGCACAGGAAAAAGCAAGTCTTGC	2152
2067	Qy	AAGACTACTCAAAATTGAAGCGAGGCAAGCAGCTCAGAACAAAAAGAAATTTCTGAAGCAGG	2126
2153	Db	AAGACTACTCAAAATTGAAGCGAGGCAAGCAGCTCAGAACAAAAAGAAATTTCTGAAGCAGG	2212
2127	Oy	AGAGGAAGAAGGMAAAGAAAAATGAAGATAAAACAAAAAGAAATTCAGATTAACAAAAAAGA	2186



Db 2765 TGAATTTCTACAGATCCCAAGTCGTCGAGAACCGCGTCAGAGACCTGGTAGGATACA 2824  
 QY 1457 CATGGACTCTCAAGAAATGCTAGATGAGTCTGATTTGCTTAAGACCCCTACCAACTACGG 1516  
 Db 2825 CTTGGCACTCGCAAGCATGCTGGATGAGTCAGAGCTGATGGTGCAGCTTCCAGACAAGA 2884  
 QY 1517 TCCAGTAGCCCTCGCCATGATGTGAATCTAGCATCATCAGCAAGTCGACTTGTTC 1576  
 Db 2885 TCGCGCTGGACCTCGCCATCGAGCTGAACCTTACAACATCGTTAGCAAGTCGCACTCTTTC 2944  
 QY 1577 AGGGTTGTGATACACAGATGATTTATGACATGTTGCTAAGATTGAAATCGGTCTCTATT 1636  
 Db 2945 AGGGCTGTGACCGGAGATGATCTTTGACATGCTGAAGGCTTCGCTCTGTGTCTACC 3004  
 QY 1637 TGCCTGGTGATTTGTCTGCAAAAAGGGAGAAATTCGCAAGGAAATGTATATCATCAAG 1696  
 Db 3005 TGCCCAACGACTATGTGTGCAAGAGGGGAGATCGGCCGTGAGATGTATCATCATCCAG 3064  
 QY 1697 ATGGAGAAGTCCAAGTCTTGGAGGCCCTGATGGTACTAAAGTTCTGTTACTCTGAAG 1756  
 Db 3065 CAGGGCAAGTGCAGGCTTTGGCGGCCCTGATGGGAAATCTGTGCTGGTGACGCTGAAG 3124  
 QY 1757 CTGGCTCGGTTTGGAGAAATCAGCTCTTAGCAGCAGGAGGAGAAACCGTCGAACCTG 1816  
 Db 3125 CTGANTCTGTGTTGGAGAAATGAAGCTTGTGCTGTGTTGGGGCGGGAAACCGGCGCAG 3184  
 QY 1817 CCAATGTGTGGCCACGCGGTTTGCCTAATCTTTTAACTCTAGACAAAAGACCCCTCCAAG 1876  
 Db 3185 CCAAGTGTGTGGCAGCGGTTTACCAACCTCTTATCTCTGGATAAGAAGGACCTGAATG 3244  
 QY 1877 AATTTCTAGTCAATATTCAGATTCGAAAGGATCTCATGAAAGAACCCAGAGTGTCTT 1936  
 Db 3245 AGATTTTGTGTGATTTCTCTGAGTCTCAGAAGTTACTTCCGGAAGAAGCCAGCGCATGC 3304  
 QY 1937 TAAGCAGAGGCTTAAGACCGCAGAG 1963  
 Db 3305 TGAGAAGCAATTAAGCCCAAGAGG 3331

RESULT 4

ID ABZ80565 standard; cdna; 2607 BP.

XX AC ABZ80565;

XX 26-JUN-2003 (first entry)

XX Nucleotide sequence of hbeta1b.

XX OCN1; OCN2; beta1b; CNG; human;

XX olfactory cyclic nucleotide gated channel subunit; ss.

XX Homo sapiens.

XX W0200304611-A2.

XX 16-JAN-2003.

XX 08-JUL-2002; 2002MO-US021184.

XX 06-JUL-2001; 2001US-0303140P.

XX 10-DEC-2001; 2001US-0337154P.

XX (SENO-) SENOMYX INC.

XX Zoller MT, Xu H, Staszewski L, Moyer B, Pronin A, Adler JB;

XX PI Servant G, Callamaras N;

XX WPI; 2003-229406/22.

XX P-PSDB; ABR39396.

XX Novel isolated nucleotide sequences encoding human OCN1, OCN2, beta 1b

XX olfactory cyclic nucleotide gated (CNG) channel subunits, useful for

PT identifying the CNG channel activators useful for enhancing smell.  
 XX Claim 14; Page 89; 97pp; English.  
 PS The present invention relates to a nucleotide sequences encoding human  
 CC OCN1, OCN2, or beta1b olfactory cyclic nucleotide gated (CNG) channel  
 CC subunit, its action being an olfactory CNG channel activity modulator. A  
 CC host cell which expresses human OCN1, OCN2 and/or beta1b is useful in a  
 CC mammalian cell-based assay for the profiling and screening of putative  
 CC modulators of a human olfactory cyclic nucleotide gated (CNG) channel.  
 CC The method is used to identify a compound as one which particularly  
 CC modulates CNG activity based on a detectable change in fluorescence. The  
 CC test cell expresses each of the human OCN1, human OCN2 and human beta1b  
 CC subunits. A fluorescence plate reader or a voltage imaging plate are used  
 CC to monitor changes in fluorescence. The compounds that activate olfactory  
 CC CNG channel enhance smell and can be used to make foods more palatable  
 CC for individuals with attenuated olfactory function. The present sequence  
 CC represents the nucleotide sequence of hOCN1  
 XX Sequence 2607 BP; 614 A; 782 C; 719 G; 492 T; 0 U; 0 Other;  
 SQ

Query Match 23.8%; Score 577.4; DB 8; Length 2607;  
 Best Local Similarity 64.3%; Pred. No. 1.2e-116;  
 Matches 866; Conservative 0; Mismatches 481; Indels 0; Gaps 0;

QY 617 TTCCAAACAGCATAGATTTCATACAGATCGACTCTATCTCTGTGGCTCTTCTTGTCA 676  
 Db 782 TTCCCAAGACATGACCGCTGACCAACCTGATGTATGCTTATGGCTGTCTTCTGG 841  
 QY 677 CTCTTCCCTTAACTCGAACTGCTGTTTATACACTGCGCTCTTCCCATATCAAA 736  
 Db 842 TGATGCTTGGAAATGGAATCTGTTGCTGATTCCTGCTGGCTGCTCCCTACAGA 901  
 QY 737 CCGCAGACAACATACACTACTGCTTATTCGGACATCATATGTATATCATCTACTTT 796  
 Db 902 CCGCGACAACATCCACACTGCTGCTGATGATTTACCTATCGACCTCATCTACTTC 961  
 QY 797 ATGATATGCTATTTATTCAGCCAGACTCCAGTTTGTAGAGAGGAGACATAATAGTG 856  
 Db 962 TGGACATACCGTGTTCAGACACGCTGCAAGTTTGTGAGAGCGGGGACATCATACGG 1021  
 QY 857 ATTCAAAATGAGCTAAGGAAACACTACAGGACTTCTACAAAATTTTCAGTTGCTCAT 916  
 Db 1022 ACAAAGGACATGCGAAATACTACTGAACTCTCGCCCTTCAAGATGGACCTGCTCA 1081  
 QY 917 CAATAATACCATTTGATATTTGCTACCTCTTCTTTGGGTTTAAATCCAAATGTTAGCAA 976  
 Db 1082 GCCTCTGCTTGGATTTTCTCTATTTGAAAGTCGGTGTGAACCCCTCTCTCGCCTGC 1141  
 QY 977 ATAGGATGTTAAAGTACACTTCAATTTTGAATTTAATCATCACCCTAGAGTCTATAATGG 1036  
 Db 1142 CCGCTGTTTAAAGTACATGGCTTCTTCGAGTTTAAACAGCGCTGGAATCCATCTCA 1201  
 QY 1037 ACAAGCATATATCTACAGAGTTTATTCGAACAACCTGATATCTTCTTTATCTGCACA 1096  
 Db 1202 GCAAGCTTACGTGTACAGGTCATCAGGACCAAGCCTACCTTCTCTACAGCTGCAAT 1261  
 QY 1097 TTAATGCTGTGTTTATTTACTGGGCTTCAAACTATGAAGGAATTTGGCACTACTAGATGG 1156  
 Db 1262 TGAATTCCTGCTTTTACTGGGCTCGGCTATCAGGCGCTCGGCTCCACTCAGTGG 1321  
 QY 1157 TGTATGATGGGGAAGGAACGAGTATCTGAGATGTTTATTTATTTGGGAGTTCGAACCTTAA 1216  
 Db 1322 TTTACGATGGCGTGGGAAACAGTTATATTCGCTGTTACTACTTTGCTGTGAAGACCTCA 1381  
 QY 1217 TTACCATTTGGCTTCCAGAACCAAACTTTATTTGAAATTTGTTTTCACACTTTGA 1276  
 Db 1382 TCACCATCGGGGCTGCTTGACCCCAAGACACTCTTTTGAATTTGTTTCCAGCTGTA 1441  
 QY 1277 ATTTTCTTCTGGAGTTTGTGTTTCTTCCAGTTTAAATTTGGTTCAGATGAGATGTTG 1336  
 Db 1442 ATTATTTACGGGCTCTTCTTCTCTGTGATCGGACAGATCAGAGATGTTGTTAG 1501

```
QY 1337 GAGCAGCTACAGCAATCAGAACTACTTCGGCGCCTGCGATGATGATGACACCAATTCCTCTACA 1396
D 1502 GGGCGGCCACCGGGGACAGACCTACTACCGCAGCTGCTGACAGCAGCGTGAAGTACA 1561
QY 1397 TGAACAATTACTCCATTCTCTAACTTGTGCMAAGCGAGTTCGGACTTCGGTATGATATA 1456
D 1562 TGAATTTCTACAAGATCCCAAGTCGCTGCAGAACCGCGTCAAGACTCTGGTACGAGTACA 1621
QY 1457 CATGGACTCTCAAGAATGCTAGATGAGTCTGATTTGCTTAAGACCCCTACCAACTACCG 1516
D 1622 CTTGGCACTCGCAAGCATGCTGGATGAGTCAGAGCTGATGCTGACGCTTCACAGACAAG 1681
QY 1517 TCCAGTTAGCCCTCGCATATGATGTAATCTTACGATCATACGAAAGTCGACTGTTTCA 1576
D 1682 TCGCGCTGACCTCGCATCGAGCTGAGTGAACATCGTTAGCAAGTCGCACTCTTTC 1741
QY 1577 AGGGTTGTCATACACAGATGATTTATGACATGTTGCTTAAGATGAAATCGGTTCTCTATT 1636
D 1742 AGGGCTGTACCGGCAGATGATCTTTGACATGCTGAAGAGGCTTCGCTCTGTGTTCTAC 1801
QY 1637 TGCCTGGTGACTTTGTCTCAAAAAGGAGAAATTTGGCAAGGAAATGTATATCATCAAGC 1696
D 1802 TGCCCAACGACTATGTGTCAGAGAGGGGAGATCGGCCGTGAGATGATATCATATCCAGG 1861
QY 1697 ATGGAGAAGTCAAGTTCTTGGAGGCCCTGATGGTACTAAAGTTCTGTGTTACTCTGAAG 1756
D 1862 CAGGGCAAGTCAGAGTCTTGGCGCGCCCTGATGGGAAATCTGTCTGTGTGAGCGCTGAAG 1921
QY 1757 CTGGGTGCGGTGTTGGAGAAATCAGCTTCTAGCAGAGGAGGAAACCGTCGAACTG 1816
D 1922 CTGGATCTGTGTTGGAGAAATGAAGCTTGTGCTGCTTGGCGCGGGAACCGCGCACGG 1981
QY 1817 CCAATGTGTGGTGGCCCAACGGGTTTGCCAAATCTTTTAACTCTAGACAAAAGACCTCCAAG 1876
D 1982 CCAACGTGTGGCGCACGGGTTTACCAACTCTTTCATCTGTGTAAGAAGGACCTGAATG 2041
QY 1877 AAATTTAGTCATTTATCCAGATTTCTGAAGGATCTCTATGAAGAACCCAGAGTGCTTT 1936
D 2042 AGATTTGTGTGATTTATCTCTGAGTCTCAGAAGTTACTCGGAAGAAAGCCAGCGCATGC 2101
QY 1937 TAAAGCAGAAGCTTAAGACCGCAGAG 1963
D 2102 TGGAAGCAACATAAGCCCAAGGAGG 2128

RESULT 5
AAS91658
ID AAS91658 standard; cDNA; 3811 BP.
XX
AC AAS91658;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #27462.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
```

WPI; 2001-639362/73.  
P-PSDB; ABG27471.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID NO 27462; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)

Sequence 3811 BP; 812 A; 1093 C; 1102 G; 804 T; 0 U; 0 Other;

Query Match 23.6%; Score 573.6; DB 5; Length 3811;

Best Local Similarity 64.6%; Pred. No. 9.6e-116;

Matches 855; Conservative 0; Mismatches 469; Indels 0; Gaps 0;

QY 640 ACAGATCCAGCTCTATCTCTCTGGCTCTTGTCTGCTCTTGCCTATAACTTGAACCTGC 699

DB 2009 ACAGACCTGATGATGCTATGCTGCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCT 2068

QY 700 TGGTTTATACCACTGCGCTCTCTTCCCATATCAACCCGAGACCAACATACACTACTGG 759

DB 2069 TGGCTGATTCCTGCTGCGCTTCCCTTACCAGACCCCGGACCAACATCACCACTGG 2128

QY 760 CTTATTGCGGACATCATATGATATCATCTACCTTTATGATATGCTATTATTCAGTCC 819

DB 2129 CTGCTGATGGATTACCTATGCGACCTCATCTACTTCTCTGGACATCACCGTGTTCAGACA 2144

QY 820 AGACTCCAGTTTGTAAAGAGGAGAGACATAATAGTGGATTCAAAATGACCTAAGGAACAC 874

DB 2189 GCGCTGCGAGTTTGTACAGAGCGGGGACATCATACGGACAAAAGGACATCGGAATAAC 2248

QY 880 TACAGGACTTCTACAAAATTTTCAGTTGGATGTCGATCAATAATACCAATTTGATATTTGC 939

DB 2249 TACCTGAAGTCTCGCGGCTTCAAGATGAGCTGCTGACGCTCTGCGCTTGGATTTCTC 2308

QY 940 TACCTCTCTTTGGGTTTAATCCAAATGTTTAGAGCAATAGATGTTTAAAGTACACTTCA 999

DB 2309 TATTGAAAGTCGGTGTGAACCCCTCTCTCGCCTGCGCCCTGTTTAAAGTACATGCCC 2368

QY 1000 TTTTGTGAATTAATCATCATCTAGAGTCTAATAATGGACAAAGCATATATCTACAGAGTT 1059

DB 2369 TTCTTCGAGTTTAACAGCGGCTGGAATCCATCTCAGCAAAAGGCTAGCTGTACAGGGTC 2428

QY 1060 ATTCGAACAACCTGGATCTTGTGTTTATCTCGACATTAATAGCTGTGTTTATTACTGG 1119

DB 2429 ATCAGGACCACAGCTACTTCTCTACAGCTGCAATTTGAATTCCTGCTTTTATTACTGG 2488

QY 1120 GCTTCAAACTATGAAGGAATTTGGCACTACTAGATGGGTGATGATGGGAAGGAACAG 1179

DB 2489 GCATCGGCCTATCAGGGCCCTCGGCTCCACTCACTGGGTTTACGATGGCGTGGGAACAGT 2548



QY 797 ATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAAAGAGGAGACATAATAGTGG 856  
D5 2457 TGGACATCACCGTGTCCAGATCGCTCCAGTTTGTCAAAGCGGGACATCATATACAG 2516  
QY 857 ATTCAAATCAGCTAAGGAACACTCAGAGACTTCTACAAATTTTCCATGTTGGATGCGAT 916  
D5 2517 ATAAGAGGAGATGCGTAATAATTAATCTGAAGTCTCAACGATTTAAGATGGACTTGTCT 2576  
QY 917 CAATAATACCATTTGATATTTGCTACTCTCTTTTGGGTTTAAATCAATGTTTAGAGCAA 976  
D5 2577 GCCTTTGGCTTTGGATTTCTCTACTTGAAGTCTGCGGTGAACCCCTTCTTGCGCTGC 2636  
QY 977 ATAGGATGTAAGTACACTTCATTTTTTGAATTTAATCATCACCTAGAGTCTATATAGG 1036  
D5 2637 CCGCTGCGCTGAAGTACATGGCTTCTTTGAGTTTAAATACCGTCTGGAAGCATCCTCA 2696  
QY 1037 ACAAGCATATATCTACAGAGTTATTTCGAACACTGGATACTTGTCTTTTATTTCGCACA 1096  
D5 2697 GCAGGCTACGTTTACAGGTTTATCAGGACCAACCCCTACTCTGTATAGCTTGCATC 2756  
QY 1097 TTAATGCTGTGTTTATTACTGGGCTTCAAACTATGAAGAAATGGCACTACTAGATGGG 1156  
D5 2757 TCAACTCTGTCTTTACTACTGGCGTGGCCCTTCCAGGGCATCGTTCCACTCACTGGG 2816  
QY 1157 TGTATGATGGGAGGAAGAAACGAGTATCTGAGATGTTATTATTGGGAGTTCGAACTTTAA 1216  
D5 2817 TTTATGACGGCTGGGGAACAGCTACTATTCGATGCTACTACTGGCTGTGAAAACTCTCA 2876  
QY 1217 TTACCAATTCGTGGCTTCCAGAACCAAACTTTATTGAAATGTTTTCACCTCTTCA 1276  
D5 2877 TCACCATCGAGACTGCGCGACCCAGACGCTCTTTGAGATCGTCTTCCAGCTGCTGA 2936  
QY 1277 ATTTTTTTCTGAGTTTTTGTGTTCTCCAGTTTTAAATGGTCAGATGAGAGATGTGATG 1336  
D5 2937 ATTTTACAGGTGCTTTCGCTTCTCTGTGATGATTGACAGATGAGAGATGTGATG 2996  
QY 1337 GAGCAGTACAGCAATCAGACTACTTCCGGCCCTGATGATGATGATGATGATGATGATG 1396  
D5 2997 GGGCGGCCACGGCGGAGAGAGTACTACCGAGCTGATGATGATGATGATGATGATGATG 3056  
QY 1397 TGAACAATTTACTTCCATTTCTAACTTGTGCAAAAGGAGTTCGGACTTGGTATGATATA 1456  
D5 3057 TGAATCTTCAAGATCCCGAGTCTGTGCAAGACCGCGTCAAGACTGATGATGATGATGATG 3116  
QY 1457 CATGGACTCTCAAGAATGCTAGATGAGTCTGATTTGCTTAAGACCTTACCAACTACGG 1516  
D5 3117 CTTGGCACTCAAGAAGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3176  
QY 1517 TCCAGTTAGCCCTCGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1576  
D5 3177 TCGCTCTGACCTGGCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3236  
QY 1577 AGGTTGTGATACACAGATGATTTATGATGATGATGATGATGATGATGATGATGATGATG 1636  
D5 3237 AGGCTGCGACCGGAGATGATCTTCAAGCATGCTCAAGCATGCTCGCTCAGTCTGCTAC 3296  
QY 1637 TCGCTGGTCACTTTGCTCAAAAAGGAGAAATTTGGCAAGGAAATGATATCATCAAGC 1696  
D5 3297 TACCCAATGACTATGTTGCAAGAGGGGAGATTTGGCCGAGAGATGATATATATCCAGG 3356  
QY 1697 ATGGAGAGTCCAAGTTCTTGGAGGCGCTGATGATGATGATGATGATGATGATGATGATG 1756  
D5 3357 CGGGGAGGTGAGGTGCTGGGCGGCCAGATGGAAGGCTGTCTCTGGTGACACTCAAG 3416  
QY 1757 CTGGGTGCGTGTGAGAAATCAGGCTTCTAGCAGCAGGAGGAGAAACCGTCAAGCTG 1816  
D5 3417 CCGGATCGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3476  
QY 1817 CCAATGTGTGCGCCACCGGTTTGGCAATCTTTTAACTCTAGACAAAAGACCTTCCAA 1876  
D5 3477 CCAATGTGTGCGCCACCGGCTTCAACCAATCTCTTCAATCTCGGATAGAGGACTTGAATG 3536  
QY 1877 AAATCTAGTGATATTCAGATTTCTGAAGAGATCTCTCATGAAAGCCAG 1928

D5 3537 AGATTTTGGTGCATTTACCTGAAATCTCAGAACTGCTCCGGAAGAAGCCAG 3588  
RESULT 7  
ADS64556  
ID ADS64556 standard; DNA; 3236 BP.  
XX ADS64556;  
XX 16-DEC-2004 (first entry)  
XX Norway rat cyclic nucleotide-gated ion channel 4.3 (CNG4.3) DNA.  
XX cyclic nucleotide concentration; screening;  
XX cyclic nucleotide-gated ion channel 4.3; CNG4.3; ds.  
XX Rattus norvegicus.  
XX WO2004083803-A2.  
XX 30-SEP-2004.  
XX 06-MAR-2004; 2004WO-EP002317.  
XX 18-MAR-2003; 2003DE-01011769.  
XX (FARB ) BAYER HEALTHCARE AG.  
XX Wunder F;  
XX WPI; 2004-728498/71.  
XX Determining intracellular concentration of cyclic nucleotides, useful for  
XX identifying, e.g. receptor ligands and enzyme inhibitors, comprises using  
XX cells that express a cyclic nucleotide activated ion channel and a  
XX photoprotein.  
XX Example; SEQ ID NO 2; 20pp; German.  
XX The invention relates to a novel method for determining the intracellular  
XX concentration of cyclic nucleotides. The method comprises preparing a  
XX cell that expresses a CNG (cyclic nucleotide-activated) ion channel  
XX together with a photoprotein and detecting the intracellular  
XX concentration of cyclic nucleotides from the luminescent signal of the  
XX photoprotein. The method may be used to screen for receptor ligands and  
XX modulators of phosphodiesterase, guanylate cyclase or nitric oxide  
XX synthase, which would be potentially useful as pharmaceuticals. The  
XX method is suitable for high throughput screening, does not require  
XX radioactive tracers, has a high signal-to-noise ratio, is very quick,  
XX highly sensitive and inexpensive. The current sequence is that of the  
XX Norway rat cyclic nucleotide-gated ion channel 4.3 (CNG4.3) DNA of the  
XX invention.  
XX Sequence 3236 BP; 780 A; 962 C; 903 G; 591 T; 0 U; 0 Other;

Query Match 23.4%; Score 568; DB 13; Length 3236;  
Best Local Similarity 64.6%; Pred. No. 1.5e-114;  
Matches 847; Conservative 0; Mismatches 465; Indels 0; Gaps 0;  
QY 617 TTCAAAACAGCATAGATTCATACACAGATCGACTCTATCTCTGTGGCTTCTTGTGCA 676  
D5 1211 TCCCCCAGAGCATCGCCCACTGACCAACCTCATGTACATCTCTGTGGCTGTCTTCTG 1270  
QY 677 CTCCTTGCCTATAACTGGAACCTGCTTTATACCACTCGGCTCTTCCCATATCAAA 736  
D5 1271 TGCTGGCTGGAACCTGGAACCTGCTGCTGATTTCTGTGGCTGGCCCTTCCCGTACCAG 1330  
QY 737 CGCGACAGACAACATACACTACTGCGCTTTATGGGACATCATATGATATCATCTACCTTT 796  
D5 1331 GGGCAGACAACATCCACCTCTGGCTGCTCATGGACTACTTGTGGCACTTCACTACCTCC 1390  
QY 797 ATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAAAGAGGAGGAGACATAATAGTGG 856

```
Db 1391 TGGACATCACCGTGTCCAGATCGCTCCAGTTTGTCAAGCGGGGACATCATTTACAG 1450
Qy 857 ATTCAATGAGCTAAGGAAACACTACAGAGCTTCTACAAATTTTCTAGTTGGATGCGCAT 916
Db 1451 ATAAGAGGAGATGCGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1510
Qy 917 CAATAATACCAATTTGATATTTGCTACTCTCTTTTGGGTTTAAATCAATTTTGTAGACAA 976
Db 1511 GCCTTTTGGCTTTTCTCTACTTGAACCTGGCGGTGAACCCCTCTTTCGCGCTGC 1570
Qy 977 ATAGGATGTTAAGTACACTTCAATTTTGAATTTAATCATCACTAGAGTCTATAATGG 1036
Db 1571 CCGCTGCTGAAATGATCATGGCTTCTTTGAGTTTAATTAACCTGCGAAGCCATCCTCA 1630
Qy 1037 ACAAGCATATATCTACAGAGTTATTCGAACAACTGATACTTGTCTTTATTTCTGCACA 1096
Db 1631 GCAAGCTACGTTTACAGGTTTATCAGGACCAACCGCTACCTGCTGTATAGCTTGATC 1690
Qy 1097 TTAATGCTGTGTTTATTAATCTGGGCTTCAAACTATGAAGGAATTTGGCACTACTAGATGG 1156
Db 1691 TCAACTCTGCTCTTACTACTGGGCTGCGCTTCCAGGGCATCGGTTCCACTCACTGGG 1750
Qy 1157 TGTATGATGGGGAAGGAAACGATCTAGATGTTTATTTATTTGGGCACTTGAACCTTTAA 1216
Db 1751 TTTATGACGGCGTGGGAAACAGCTACATTCGATCTACTACTGGGCTGTGAAAACCTCTCA 1810
Qy 1217 TTACCATTTGCTGGCTTCCAGAACCAAACTTTATTTGAAATTTGTTTCAACTCTTGA 1276
Db 1811 TCACCATCGGAGACTGCCGACCCCGACGCTCTTTGAGATGCTCTCCAGCTGCTGA 1870
Qy 1277 ATTTTCTTCTGGAGTTTTTGTGTTCTCCAGTTTAAATTTGGTCAGATGAGAGATGTGATTG 1336
Db 1871 ATTATTTTACAGGTGCTCTGCTTCTCTGTGATGATTGACAGATGAGAGATGTGTTGG 1930
Qy 1337 GAGAGCTACAGCAATCAGAACTACTTCCGCGCTGATGATGATGATGATGATGATGATGATG 1396
Db 1931 GGGCGCGCACGGGAGGACAGCTACTTACCAGCTGATGATGATGATGATGATGATGATGATG 1990
Qy 1397 TGAACAAATTAATCTCAATCTTAACTTGTGCAAAAGCGAGTTCCGACTTGTGATGAATATA 1456
Db 1991 TGAATTTCTAAGATCCCGAGGTCTGTGCAAGAACCCGCTCAAGACCTGTTATGATGATCA 2050
Qy 1457 CATGGGACTCTCAAGAAATGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1516
Db 2051 CTTGGCAATCACAAGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2110
Qy 1517 TCCAGTTAGCCCTCGCCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1576
Db 2111 TGGCTCTGGACCTGGCCATTTGACGTAAACTACAACTTGTGATGATGATGATGATGATGATG 2170
Qy 1577 AGGCTGTGTATACACAGATGATTTATGACATGTTTGTGATGATGATGATGATGATGATGATG 1636
Db 2171 AGGCTGTGTATACACAGATGATTTATGACATGTTTGTGATGATGATGATGATGATGATGATG 2230
Qy 1637 TGCCTGTGTGATTTTGTCTGCAAAAGGAGAGAAATTTGGCAAGAAATGATGATGATGATGATG 1696
Db 2231 TACCAATGATGATGTTGTGCAAGAGGGGAGATTTGGCCGAGAGATGATGATGATGATGATG 2290
Qy 1697 ATGAGAGATGCAAGTTCTTGGAGCCCTGATGATGATGATGATGATGATGATGATGATGATG 1756
Db 2291 CGGGCGAGGTGCAGTGTGCGGCGCCGACAGATGAAAGGCTGTCTCTGTGATGATGATGATG 2350
Qy 1757 CTGGGTGGTGTGGAGAAATCAGCTTCTAGCAGCAGGAGGAGGAAACCGTTCGAATCTG 1816
Db 2351 CCGATCGGTGTTTGGAGAGATAGCTTGTGCTGTGCGGGGGGGTAAACAGGCGCAGG 2410
Qy 1817 CCAATGTGGTGGCCACCGGTTTGGCAATCTTTTAACTCTAGACAAAAGACCTCTCCAAAG 1876
Db 2411 CCAATGTGGTGGCCACCGGTTTCCAACTCTTCAATCTCTTCAATCTGATGATGATGATGATG 2470
Qy 1877 AAATTTAGTGCATTTATCCAGATTTCTGAAAGGATCTCTCATGAAAGAGCCAG 1928
```

```
Db 2471 AGATTTTGTGTCATTACCTGTAATCTCAGAAGCTGCTCCGGAAGAAGGCCAG 2522
RESULT 8
AAS76156
ID AAS76156 standard; cDNA; 2244 BP.
XX AAS76156;
AC AAS76156;
XX 13-FEB-2002 (first entry)
DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #11960.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
OS Homo sapiens.
XX WO200175067-A2.
PN 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
DR P-PSDB; ABG11969.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 11960; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 2244 BP; 720 A; 469 C; 592 G; 463 T; 0 U; 0 Other;
Query Match 21.3%; Score 518; DB 5; Length 2244;
Best Local Similarity 94.7%; Pred. No. 1.3e-103;
Matches 536; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 287 GAACAGTGCACAGCAGAGAAATGAGCCCGGAAAGAGGTCCAAACAGCCCAAAA 346
Db 83 GCMAACTATCACAGGACAGAAACCAACACCCGATGTTCTTCACACAGCCCAAAA 142
```

QY 347 ACAAACCGCTCGAGCTCTCTGTATATAATGAGTATGCCGATGCCAGCTACACAACCTGG 406  
Db |||||  
QY 143 ACAAAACCGCTCGAGCTCTCTGTATATAATGAGTATGCCGATGCCAGCTACACAACCTGG 202  
Db |||||  
QY 407 TGAAGAAGTATGGTCAAGAACAGCCCTCTACAAAGAAAAGTTGGTAGAGGGAGATCTCT 466  
Db |||||  
QY 203 TGAAGAAGTATGGTCAAGAACAGCCCTCTACAAAGAAAAGTTGGTAGAGGGAGATCTCT 262  
Db |||||  
QY 467 CCTCACCCGAAGCCAGCCACAAACTGCAAAAGCCACACGGCTGTACCAACAGTAAAGAAA 526  
Db |||||  
QY 263 CCTCACCCGAAGCCAGCCACAAACTGCAAAAGCCACACGGCTGTACCAACAGTAAAGAAA 322  
Db |||||  
QY 527 GCGATGATAAGCAACAGAACATTAACAAGCTGTGTGGTTCAAAGTCAAAAAGATCC 586  
Db |||||  
QY 323 GCGATGATAAGCAACAGAACATTAACAAGCTGTGTGGTTCAAAGTCAAAAAGATCC 382  
Db |||||  
QY 587 CTTTAAACAGTACTTAAAGCGAATTAACCTTCCAAACAGCATAGATTCATACACAGATC 646  
Db |||||  
QY 383 CTTTAAACAGTACTTAAAGCGAATTAACCTTCCAAACAGCATAGATTCATACACAGATC 442  
Db |||||  
QY 647 GACTCTATCTCTGTGGCTCTTGTCTGTGTCACCTCTTGCCTATAACTGGAATCTGTGTTTA 706  
Db |||||  
QY 443 GACTCTATCTCTGTGGCTCTTGTCTGTGTCACCTCTTGCCTATAACTGGAATCTGTGTTTA 502  
Db |||||  
QY 707 TACCACCTGCGCTCTGCTTCCCATATCAAAACCGCAGACAACATACACTACTGGCTTATTG 766  
Db |||||  
QY 503 TACCACCTGCGCTCTGCTTCCCATATCAAAACCGCAGACAACATACACTACTGGCTTATTG 562  
Db |||||  
QY 767 CCGACATCATATGTGATATCATCTACCTTTATGATATGCTATTATCCAGCCAGACTCC 826  
Db |||||  
QY 563 CCGACATCATATGTGATATCATCTACCTTTATGATATGCTATTATCCAGCCAGACTCC 622  
Db |||||  
QY 827 AGTTTGTAAAGAGGAGGAGACATAATA 852  
Db |||||  
QY 623 AGTTTGTAAAGAGGAGGAGACATAATA 648  
Db |||||

RESULT 9

AA576184/c  
ID AA576184 standard; cdNA; 2244 BP.  
XX  
AC AA576184;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #11988.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR P-PSDB; ABG11997.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
XX biodiversity.

PS Claim 1; SEQ ID NO 11988; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AA564197-AA594564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

Sequence 2244 BP; 463 A; 592 C; 469 G; 720 T; 0 U; 0 Other;

Query Match 21.3%; Score 518; DB 5; Length 2244;  
Best Local Similarity 94.7%; Pred. No. 1.3e-103;  
Matches 536; Conservative 0; Mismatches 30; Indels 0; Gaps 0;  
QY 287 GAACAGTCCGAGCAGAGGAAATGGACCCCGGAAAGAGGTCCAAACAGCTCCACAA 146  
Db |||||  
QY 347 ACAACCCGCTCGAGCTCTCTGTATATAATGAGTATGCCGATGCCAGCTACACAACCTGG 406  
Db |||||  
QY 2102 ACAACCCGCTCGAGCTCTCTGTATATAATGAGTATGCCGATGCCAGCTACACAACCTGG; 204  
QY 407 TGAAGAAGTATGGTCAAGAACAGCCCTCTACAGAAAAGTTGGTAGAGGGAGATCTCT 466  
Db |||||  
QY 2042 TGAAGAAGTATGGTCAAGAACAGCCCTCTACAGAAAAGTTGGTAGAGGGAGATCTCT 198  
QY 467 CCTCACCCGAAGCCAGCCACAAACTGCAAGCCACACGGCTGTACCAACAGTAAAGAAA 526  
Db |||||  
QY 1982 CCTCACCCGAAGCCAGCCACAAACTGCAAGCCACACGGCTGTACCAACAGTAAAGAAA 1923  
QY 527 GCGATGATAAGCCACAGAACATTAACAAGCTGTGTGGTTCAAAGTCAAAAAGATCC 586  
Db |||||  
QY 1922 GCGATGATAAGCCACAGAACATTAACAAGCTGTGTGGTTCAAAGTCAAAAAGATCC 1863  
QY 587 CTTTAAACAGTACTTAAAGCGAATTAACCTTCCAAACAGCATAGATTCATACACAGATC 646  
Db |||||  
QY 1862 CTTTAAACAGTACTTAAAGCGAATTAACCTTCCAAACAGCATAGATTCATACACAGATC 1803  
QY 647 GACTCTATCTCTGTGGCTCTTGTGTCACCTCTTGCCTATAACTGGAATCTGTGTTTA 706  
Db |||||  
QY 1802 GACTCTATCTCTGTGGCTCTTGTGTCACCTCTTGCCTATAACTGGAATCTGTGTTTA 1743  
QY 707 TACCACCTGCGCTCTGCTTCCCATATCAAAACCGCAGACAACATACACTACTGGCTTATTG 766  
Db |||||  
QY 1742 TACCACCTGCGCTCTGCTTCCCATATCAAAACCGCAGACAACATACACTACTGGCTTATTG 1683  
QY 767 CCGACATCATATGTGATATCATCTACCTTTATGATATGCTATTATCCAGCCAGACTCC 826  
Db |||||  
QY 1682 CCGACATCATATGTGATATCATCTACCTTTATGATATGCTATTATCCAGCCAGACTCC 1623  
QY 827 AGTTTGTAAAGAGGAGGAGACATAATA 852  
Db |||||  
QY 1622 AGTTTGTAAAGAGGAGGAGACATAATA 1597  
Db |||||

RESULT 10  
ABU19457

ID ABL19457 standard; DNA; 3275 BP.  
XX ABL19457;  
AC  
XX  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 9844.  
DE  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW  
KW pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
PF  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR  
PR 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX  
XX WPI; 2001-656860/75.  
DR  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
XX  
XX Claim 1; SEQ ID NO 9844; 21pp + Sequence Listing; English.  
PS  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX  
SQ Sequence 3275 BP; 843 A; 849 C; 873 G; 710 T; 0 U; 0 Other;

Query Match 9.28; Score 223; DB 4; Length 3275;  
Best Local Similarity 48.74; Pred. No. 1.1e-38;  
Matches 635; Conservative 0; Mismatches 665; Indels 3; Gaps 1;  
QY 650 TCTATCTCCGTGGCTCTGCTGTCTCACTCTGCTATTAAGTGAAGTCTGGTTTATAC 709  
DB 1552 TCTATCTCTGGCTATGGTGTATCGTATCGTCTCTATTAAGCTCTGGTATCC 1611  
QY 710 CACTGCGCCTCGTCTCCCATATCAAAACCGCAGACAACTACTGCTTATTGCGG 769  
DB 1612 CGCTGCGCTCGTCTTCCCTTTCAAAACCAAGGAGAACCAATATCTGGCTGGCTGG 1671  
QY 770 ACATCATATGTGATATCATCTACCTTTATGATATGCTATTTATCCAGCCAGCTCCAGT 829  
DB 1672 ACTTTTGGCTGTGACATCATATATCTACTGGATGGTCTCTTTTAAACATCGGGTTATGT 1731  
QY 830 TTGTAAGAGGAGGAGACATATAGTGATTCAAATGAGCTAAAGGAAACACTACAGGACTT 889  
DB 1732 ACCTTTTGGAGGATCTGGGTGAAACCAAGAACCTCACCAAGAAAGACTACATGCGAA 1791  
QY 890 CTACAAAATTTAGTGGATGTCGATCAATTAATACCAATTTGATTTGCTTCTTCT 949  
DB 1792 AACTGCAAGTTTAAAGCTGATCTGCGCCCTCTTCCGCTGGAGCTATGTACTTTAAAC 1851  
QY 950 TTGGGTTTAAATCCAAAT---GTTTAAAGCAATAGATGTTAAAGTACACTTTTGTG 1006

DB 1852 TCAGAACACAGCGGTTGGTTGGCTTTTCCAAAGGTTCTTTAAGATTCCAGAGCTTCTGGG 1911  
QY 1007 AATTTAATCATCACCTAGAGTCTATAATAGGACAAAGCATATATCTACAGAGTTATTGAA 1066  
DB 1912 AAGTGTTCGCTCTTTGGATCGAGTCATATCTTCCGACACTTTGTTCAGTGGCCAAA 1971  
QY 1067 CAACCTGGATACCTTGTGTTTATCTTGCACATTAATGCTGTGTTTATTTACTGGCTTCAA 1126  
DB 1972 CCTAACCTACATGCTCTACATGATCCACATCAGCGTGTCTGTACTACGCTATAGTG 2031  
QY 1127 ACTATGAAGGAATTTGGCACTACTAGATGGGTGTATGATGGGAAGGAAACAGAGTATCTCA 1186  
DB 2032 ACTACAGGGTCTGGGACAGAAATCGTTGGTCTTCAGCGGCAAGGGCATCCCTATGTAA 2091  
QY 1187 GATGTTATTATTGGGCAAGTTCCAACTTTTAATTTACCATTTGGTGGCTTCCAGAACCAAA 1246  
DB 2092 GGTGCTTTGCTTTTGGCCACCAAGACGGCAACTTCGATAGGAAAAAACCACAGCCGAAC 2151  
QY 1247 CTTTATTTGAAATTTGTTTTTCAACTCTTCACTTTTCTGGAGTTTCTTGTCTTCCA 1306  
DB 2152 GTCAGGAGAGTACGCTCTTCATGACGGTGGCTGGCTAATGGCGTCTTTGTTGCTGCTC 2211  
QY 1307 GTTTAATTTGGTTCAGATGAGAGATGTGATTGGAGCAGCTACAGCCAAATACAGAACTATCTC 1366  
DB 2212 TGCCTTATCGTTCAGATCAGGATATCATCTCGACAGCCACTCGGAATAAGCATGAGTATC 2271  
QY 1367 GCGCTGTCATGATGACACCATTTGCTTACATGAACAATTTACTCCATCTCTAAACITTTGTC 1426  
DB 2272 GCGAGCTGGAGATGAGACACTCGAGTACATCGACGCTCAATCTCTCCAGGAGGTGC 2331  
QY 1427 AAAAGCGACTTCGGACTTGGTATGATATACATGGGACTCTCAAAAGAAATGCTAGATGAGT 1486  
DB 2332 AGTCGAGGCTCAAAATGTGGTTCAGTTTACATGGGAGCAGCAGCCACTTTGGATGAAT 2391  
QY 1487 CTGATTGTCTTAAGACCTTACCAACTACCGTTCAGTTAGCCCTTCGCTTGGATGTGAAT 1546  
DB 2392 CCACATATTTGGATGCTTTTGGCCATCAATTTGAAACCCGATATCGCCATCTCCGTGACA 2451  
QY 1547 TCAGCATCATACGAAAGTCGACTTGTTCAGAGGTTGTGATACAGATGATTTATGACA 1606  
DB 2452 TCCAAACGCTCTCCAAAGTGCAACTCTTTGCTGATTTGCGAGGAGGCTTTACTCAGAGATC 2511  
QY 1607 TGTGTTGTAAGATTGAATTCGGTCTCTATTTGCTGCTGACTTTGCTGCAAAAAGGGAG 1666  
DB 2512 TGGTCTTAAACTAAGGCGAGTCACATTTTTCGAGGAGATTTTCGTTTGGCCGAAAGCGC 2571  
QY 1667 AATTTGGCAAGGAATGTATATCATCAAGCATGGAAGTCCAAAGTTCTTTGGAGGCCCTG 1726  
DB 2572 AAGTGGCAGAGAGATGTATATCGTGAACCTGGACAGGTTCAAGTATGATGGTGGTCTTA 2631  
QY 1727 ATGGTACTAAAGTTCTGGTTACTCTCTAAAGCTGGGTGGTGGTGGAGAAATCAGCCCTTC 1786  
DB 2632 GTAGCGATGTAGTGTGGCTGCTTAAACGAAAGTTCCGTTTTCGGGGAGATCAGTCTGC 2691  
QY 1787 TAGCAGGAGGAGAGAAACCGTCGAACTGCCAATGTGGTGGCCAGGGGTTTGGCAATC 1846  
DB 2692 TGGGAATAAATGGAGCGGATTCGAGAACAGCAGATGTCCGATCCAAAGGGCTACTTCCAAC 2751  
QY 1847 TTTTAACTCTAGACAAAAGAACCCCTCCAAAGAAATTTCTAGTGCATTTATCCAGATTTCTGAAA 1906  
DB 2752 TGTTCGTGCTCCAAATCGGATTTAAACGAGGTGCTTGGCTACTATCCACGCCCCAGG 2811  
QY 1907 GGATCCTCATGAAGAAAGCAGAGTGTCTTTTAAAGCAGAAAGC 1949  
DB 2812 CGATTCTCAAAAAGCGTGTCTCGTCAGCTAATGCGAAAAAATGC 2854

RESULT 11  
ADD93243  
ID ADD93243 standard; cDNA; 2500 BP.  
XX  
XX  
AC ADD93243;  
XX

DT	29-JAN-2004	(first entry)	
DE	Cyclic nucleotide gated channel alpha 1 coding sequence.		
XX	ss; gene;	cyclic nucleotide gated channel alpha 1 ;	
KW	central nervous system;	CNS; eye; blood-brain barrier;	
KW	blood-retina barrier;	angiogenesis ; revascularisation;	
KW	retinal pigment epithelium;	RPE; neurosensory retina ; choroida ;	
KW	age-related macular degeneration;	AMD; diabetic retinopathy;	
KW	Alzheimer's disease;	Parkinson's disease; depression; bipolar disorder;	
KW	schizophrenia; amnesia;	headache; stroke; insomnia; alcohol abuse;	
KW	anxiety; chronic pain;	retinal blastoma; primary retinal detachment .	
XX	Homo sapiens.		
OS			
XX	Key	Location/Qualifiers	
XX	CDS	25..2097	
FT		/tags= a	
FT		/product= "Cyclic nucleotide gated channel alpha 1"	
XX	WO2003087368-A2.		
XX	23-OCT-2003.		
XX	16-APR-2003;	2003WO-EP004003.	
XX	18-APR-2002;	2002EP-00008761.	
PR	05-DEC-2002;	2002US-0431173P.	
XX	(LYNK-) LYNKEUS BIO TECH GMBH.		
XX	Drumm K, Schloer SH, Goehring F;		
PI	WPI; 2003-845328/78.		
DR	P-PSDB; ADD93244.		
XX	Treating central nervous system and/or eye disorders comprises		
PT	administering to a subject a composition comprising a compound that		
PT	modulates a target gene or gene product, outside the blood-brain and/or		
PT	blood-retina barriers.		
XX	Claim 19; SEQ ID NO 1; 91pp; English.		
PS			
XX	This sequence encodes the human cyclic nucleotide gated channel alpha 1		
CC	protein. This cDNA may be targeted in the method of the invention for		
CC	treating a disorder of the central nervous system (CNS) and/or the eye.		
CC	The method comprises administering to a subject a composition comprising		
CC	a compound capable of modulating a target gene or gene product in a		
CC	therapeutic amount, where the composition is administered outside the		
CC	blood-brain and/or the blood-retina barriers. The composition is in a		
CC	designed form to be introduced into the cells or tissue of the CNS or eye		
CC	by a suitable carrier, characterized by the application occurring outside		
CC	the blood-brain or blood-retina barriers. It is also designed for		
CC	systemic administration or for administration by iontophoresis, or for		
CC	retrobulbar application or as eye drops. The compound, nucleic acid		
CC	molecule or vector is useful for the preparation of a composition for		
CC	treating, preventing and/or delaying a disorder of the CNS and/or the		
CC	eye, e.g. related to angiogenesis and/or neovascularisation, retinal		
CC	pigment epithelium (RPE), neurosensory retina and/or choroida and wet		
CC	age-related macular degeneration (AMD) or diabetic retinopathy. The		
CC	compound may also be used as a lead compound in drug discovery and		
CC	preparation of drugs or prodrugs. The nucleic acid molecule and		
CC	polypeptide are also used for the validation of test substances, lead		
CC	compounds, drugs and prodrugs for the treatment of the disorder mentioned		
CC	above or for the identification and isolation of downstream genes, which		
CC	respond to modulation of a gene comprising the nucleic acid molecule or		
CC	its encoded gene product. The method is used for diagnosing, preventing		
CC	or treating the disorder cited above, which may include Alzheimer's		
CC	disease, Parkinson's disease, depression, bipolar disorder,		
CC	schizophrenia, amnesia, headache, stroke, insomnia, alcohol abuse,		
CC	anxiety, chronic pain, retinal blastoma, primary retinal detachment or		
CC	age-related macular degeneration.		
XX			

SQ	Sequence	2500 BP; 861 A; 431 C; 525 G; 683 T; 0 U; 0 Other;	
	Query Match	7.3%; Score 176.4; DB 10; Length 2500;	
	Best Local Similarity	49.6%; Pred No. 1.7e-28;	
	Matches	617; Conservative 0; Mismatches 591; Indels 36; Gaps 5;	
QY	740	CAGACAACATACACTACTGGCTTATTGGGACATCATATGTGATATCATCTACCTTTATG	799
DB	605	CTGATTACCTAGAAATTTGGCTCATTTTGGATTACGTATCAGACATAGTCTATTAAATCG	664
QY	800	ATATCTCTATTTATCCAGCCAGACTCCAGTTTGTGAAGGAGGAGACATAATAGTGGATT	844
DB	665	ATATG---TTTGTACGAACAAGGACAGGTACTCTAGAAACAAGGACTGCTGGTAAAGGAAG	721
QY	860	CAAAATGAGCTAAGGAAACACTACAGGACTTCTACAAAATTTTCAGTTGGATCGCATCAA	914
DB	722	ACCTTAAACTATAAATAAATAAATCCAACTTGCAATTTAAACTTGATGTTCTGTCTAC	781
QY	920	TAATACCAATTTGATATTTGCTTACCTCTCTTTGGGTTTA---ATCCAATGTTTAGAGCAA	976
DB	782	TGATACCAACTGATTTGCTGTATTTTAAGTTTAGGTGGAACCTATCCAGAAATTAGATTAA	841
QY	977	ATAGCATGTTAAAGTACACTTCATTTTGAATTTAATCATCATCCTAGAGTCTATAATGG	1036
DB	842	ACAGGTTGTTACGGTCTCTCGTATGTTTGAGTTCTTCCAGAGAACAGAAACAAGGACAA	901
QY	1037	ACAAAGCATATATCTACAGAGTTATTTCGAACAACCTGGGATACTTGTCTTTATTTCTGCACA	1096
DB	902	ACTATCCAAACATCTTCAGGATTTCCAACTTGTATGTATATCGTCATCATTTACCACT	961
QY	1097	TTAATGCTGTGTTTATTACTGGGCTTCAAACTATGAAGGAAATGGGCACTACTAGATGGG	1156
DB	962	GGATGCACTGTGTTCTACTCTATTTCTAAAGCTATTGGGATTTGGAATGATACATCGG	1021
QY	1157	TGTA-----TGATGGGAAGGAACGAGTATCTGAGATGTT	1192
DB	1022	TCTACCTCGATATTAAATGATCCTGAATTTGGCGTTTGGCTAGAAAATACGTATACAGCC	1081
QY	1193	ATTATTGGGCACTTCGAACCTTTAATTTACATTTGGTGGCTTCCAGAACACACAACTTTAT	1252
DB	1082	TTTACTGCTACACTGACTTTGACTACATTTGGTGAACACCCCTCCGCTGAGGGATT	1141
QY	1253	TGAAATTTGTTTTCAACTCTTGAAATTTTTTTTGGAGTTTTTGTGTCTCCAGTTAA	1312
DB	1142	CTGATATGCTCTTTGTGGTGGTGTATTTCTTAATTTGGAGTGTAAATTTTGTCTACCATCG	1201
QY	1313	TTGGTCAGATGAGATGTTGAGGAGCTACAGCAATCAGNACTACTTCCCGCTT	1372
DB	1202	TGGTAAACATAGGTTCTATGATTTTCCAAACATGAATGCGAGCCAGACAGAAATTTCAAG	1261
QY	1373	GCATGGATGACACCACTTGCCTACATGAACAATTTACTCTTAACTCTGTGCAAAAT	1432
DB	1262	GAAATGATGCTATCAAGCAATATATGCAATTTTCGAAATGTAAGCAAGATATGCAAAAGA	1371
QY	1433	GAGTTGGAATTTGGTATGAATATACATGGGACTCTCAAAAGAAATGCTAGATGAGTCTGAT	1492
DB	1322	GGGTTATTAAATTTGGTTTGACTACCTGTGGACCAACAAAAAACAAGTTGATGAGAAAGA	1381
QY	1493	TGCTTAAGACCCCTACCACTACGGTCCAGTTTAGCCCTCGCCATTTGATGTGAATTCAGCA	1552
DB	1382	TCTTAAAGTATCTACCTGATAAACCTAAGAGCAGAAATTTGCCATCAACCTTCACTTAGACA	1441
QY	1553	TCATCAGCAAAAGTGCAGCTTGTTCAGGGTGTGTATACACATGATTTTATGACATGTTGC	1612
DB	1442	CATTAAAAAAGGTACCGCATTTTTCGTGATTTGGAAGCTGGTCTGTGTGGTGGATTTGCT	1501
QY	1613	TAAGATTGAATTCGGTCTCTCTATTGCTGTGTGACTTTTGTCTGCAAAAAGGAGAAATG	1672
DB	1502	TGAATTTGCAACCCCAAGTCTACAGTCTCGAGATTAATTTTGAAGAAGAGGGGATATCG	1561
QY	1673	GCAAGGAATCTATATCATCAAGCATGGAGAGTCCCAAGTTCTTTGGAGGCCCTGATGTA	1732
DB	1562	GACGAGATGTATACATTTATCAAGGAAGGCAAACTCGCTGTGGTGGCAGATGATGAGTCA	1621

QY 1733 CTAAAGTTCTGGT-----TACTCTGAAGAGCTGGTGGTGGTGGAGAAATC--AGCCTTC 1786  
 DB 1622 CTCAGTTTCTGGTATTGAGCGATGGCAGCACTTCGGTGGATCAGCATTTCTTAACATTA 1681  
 QY 1787 TAGCAGCAGGAGGAGAAACCGTCGAACCTGCAATGTGGTGGCCACGGGTTGGCAATC 1846  
 DB 1682 AAGGAGCAAGCTGGCAATCGNAGAACCGCCNATATTAAAGTAGTATGGCTACTGAGCC 1741  
 QY 1847 TTTTAACTCTAGACAAAAGACCTCCAAAGAAATTTCTAGTGGCAATATCCAGATTTCTGAAA 1906  
 DB 1742 TGTCTGTCTCTCAAAAGATGACCTCATGAAGCTCTAACTGAGTAGTACCAGATGCCAAA 1801  
 QY 1907 GGAATCCTCATGAAGAAAGCAGAGTCTTTTAAAGCAGAGGCT 1950  
 DB 1802 CTATGCTAGAAAGAAAGGAAGCAAAATTTTATGAAGATGGT 1845

RESULT 12

AD24566  
 ID AD24566 standard; DNA; 2500 BP.

AC AD24566;

DT 21-OCT-2004 (first entry)

XX Breast cancer prognosis marker #427.

DE db; breast cancer; prognosis; gene expression; diagnosis.

KW Homo sapiens.

OS WO2004065545-A2.

PN 05-AUG-2004.

XX 15-JAN-2004; 2004WO-US001100.

XX 15-JAN-2003; 2003US-00342887.

PA (ROSE-) ROSETTA INPHARMATICS LLC.

PA (NECA-) NETHERLANDS CANCER INST.

XX Van't Veer LJ, He Y;

XX WPI; 2004-593473/57.

XX Classifying a breast cancer patient according to prognosis comprises  
 PT determining the similarity between the level of expression of each of  
 PT five genes in a cell sample taken from patient, to control levels.

XX Disclosure; SEQ ID NO 427; 226pp; English.

XX The invention relates to a method of classifying a breast cancer patient  
 CC according to prognosis by determining the similarity between the level of  
 CC expression of each of five genes for which markers are listed in the  
 CC specification, in a cell sample taken from the breast cancer patient, to  
 CC control levels of expression for each respective five genes to obtain a  
 CC patient similarity value. The methods are useful for classifying a breast  
 CC cancer patient according to prognosis. Kits and computer program products  
 CC are useful for data analysis using the diagnostic, prognostic and  
 CC statistical methods of the invention. This sequence corresponds to a  
 CC marker used in the method of the invention.

SQ Sequence 2500 BP; 861 A; 431 C; 525 G; 683 T; 0 U; 0 Other;

Query Match 7.3%; Score 176.4; DB 13; Length 2500;

Best Local Similarity 49.6%; Pred. No. 1.7e-28;

Matches 617; Conservative 0; Mismatches 591; Indels 36; Gaps 5;

QY 740 CAGACACATACACTACTGCTTATGGGACATCATATGTGATATCATCTACCTTTATG 799

DB 605 CTGATTACCTAGAAATATGGCTCAATTTGGATTACGATCAGACATAGTCTTATTAATCG 664

QY 800 ATATGCTATTTATCAGCCAGACTCCAGTTTGTAAAGAGGAGACATATAAGTGGATT 859  
 DB 665 ATATG---TTTGATCAACAAGGACAGGTACTAGAACAGGACTGCTGGTAAAGGAAG 721  
 QY 860 CAAATGAGCTAAGGAAACACTACAGGACTTCTACAAAATTTTCAGTTGGATGTCGATCAA 919  
 DB 722 AACTTAAACTATAATAATAATAATCAACTTGCATTTTAAACTTGGTCTGCTCAC 781  
 QY 920 TAAATACCTTTGATATTTGCTACCTCTCTTTGGGGTTTA---ATCAATGTTTAGAGCAA 976  
 DB 782 TGATACCAACTGATTTGCTGTATTTTAAGTTAGGGTGGAACTATCCAGAAATTAGATTAA 841  
 QY 977 ATAGGATGTTAAAGTACACTTCAATTTTGAATTTTAAATCATCACCTAGAGTCTATATCG 1036  
 DB 842 ACAGGTTGTTACGGTCTCTCGTATGTTTGAAGTCTTCCAGAGAACAGAAACAGGACAA 901  
 QY 1037 ACAAGCATATATCTACAGAGTTTATTCGAAACACTGGATACTTGTCTGTTTATTTCTGCACA 1096  
 DB 902 ACTATCCAAACATCTTCAGGATTTCCAACTTGTATGTATATCGTCATCATTTATCCACT 961  
 QY 1097 TTAATGCCCTGTGTTTATTACTGGGTTTCAAACCTATGAAGAAATTTGGCACTACTAGATGG 1156  
 DB 962 GGAATGCTGTGTGTTCTACTCTATTTCTAAAGCTATTTCGATTTGGAATGATACATGGG 1021  
 QY 1157 TGTA-----TGATGGGAAGAAACGAGTATCTGAGATGTT 1192  
 DB 1022 TCTACCCCTGATATTAATGATCTCGAATTTGGCGTTTGGCTAGAAAATACGTATACAGCC 1081  
 QY 1193 ATTATTTGGGCGAGTTTGAACCTTTAAATTTACCATTTGGTGGCTTCCAGAAACACAAACTTTAT 1252  
 DB 1082 TTTACTGGTCTACATGACTTTTGAATTTGATACCATTTGGTGAACACCCCTCCGTCGAGGAT 1141  
 QY 1253 TTGAAATGTTTTTCAACTCTTGAATTTTTTTTCTGAGATTTTGTGTTCTCCAGTTTAA 1312  
 DB 1142 CTGAGTATGTTCTTTGGTGGTGGTGAATTTCTTAATGGAGTGTAAATTTTGTACCATCG 1201  
 QY 1313 TTGGTCAGATGAGAGATGTGATTTGAGCAGCTACAGCCCAATCAGAACTACTTCCGGCCCT 1372  
 DB 1202 TTGGTAAACATAGGTTCTATGATTTCCAAACATGAATCGACCCAGAGAGAGAAATTTCAAGCAA 1361  
 QY 1373 GCATGATGACACACCATTTGCCTACATCAACAAATTTACTCCATTTCTCAAAATTTGTCAAAAGC 1432  
 DB 1262 GAATGCTCTATCAAGCAATATATGCATTTTCCAAATGTAAAGCAAGATATGCAAAAGA 1321  
 QY 1433 GAGTTCGGACTTTGGTATGAATATACATGGGACTTCTCAAGAAATGCTAGATGAGTCTGATT 1492  
 DB 1322 GGGTTATTAAATGGTTTGTACTACCTGTGCACCAACAAACAAACAAACAGTTGATGAGAAAGAG 1381  
 QY 1493 TGCTTAAAGACCTTACCACCTACGGTCCAGTTCAGCTTACCCCTCGCCATTGATGTGAACCTCAGCA 1552  
 DB 1382 TCTTAAAGTATCTACCTGATAAACTAAGAGCAGAAATTTGCCATCAACGTTTCACTTACAGCA 1441  
 QY 1553 TCATCAGCAAGTCGACTTTGTTCAAGGTTTGTATACACAGATGATTATGACATGTTGC 1612  
 DB 1442 CATTAAGAAAGGTACGCAATTTTGTGATTTGCAAGCTGGTCTGTTGGTGGAGTTGGTCT 1501  
 QY 1613 TAAGATTGAAATCCGTTCTCTATTTCCTCGTGGTGAATTTGTCTCAAAAGGAGGAAATTTG 1672  
 DB 1502 TGAATTTGCAACCCCAAGTCTACAGTCTCGAGATTATATTTGCAAGAAAGGGGATATCG 1561  
 QY 1673 GCAAGGAAATGTATATCATCAAGCATGGAGAACTCCAAAGTCTTTGAGAGCCCTTGATGGTA 1732  
 DB 1562 GAGCAGAGATGTACATTATCAAGGAAGGCAAACTCGCTGTGGTGGCAGATGATGGAGTCA 1621  
 QY 1733 CTAAAGTTCTGGT-----TACTCTGAAGCTGGGTCGGTGGTGGAGAAATC--AGCCTTC 1786  
 DB 1622 CTCAGTTTGGTGTATTGAGCGATGGCAGCACTTCCTGGTGGATCAGCATTTCTTAACATTA 1681  
 QY 1787 TAGCAGCAGGAGGAGGAAACCGTCGAATCTGCCAATGTGGTGGCCACCGGTTTGGCAATC 1846  
 DB 1682 AAGGAGCAAGCTGGCAATCGAAGAAACCGCCNATATTAAAGTAGTATGGCTACTGAGACC 1741

DB	143	ACAAACCGGCTGCAGCTCTCTGTTATATAATGAGTATGCCGATGCCAGCTACACAACTCG	202
Qy	407	TGAAAAAGATGCGTCAAAAGAACAGCCCTCTACAAAGAAAAAGTTGGTAGGGAGATCTCT	466
Db	203	TGAAAAGATGCGTCAAGAACAGCCCTCTACAAAGAAAAAGTTGGTAGGGAGATCTCT	262
Qy	467	CCTCACCGGAGCCAGCCCAAAACTG	493
Db	263	CCTCACCGGAAGCCAGCCCAAACTG	289

RESULT 14  
ABL18085  
ID ABL18085 standard; DNA: 2109 BP.

ID	ABL18085 standard; DNA; 2109 Sp.
XX	
AC	ABL18085;

XX	ABL18085;
AC	
DT	26-MAR-2002 (first entry)

XX 26-MAR-2002 (first entry)  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 5728.

DE *Drosophila melanogaster* genomic polynucleotide SEQ ID NO 5728.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;

KW prokaryota; developmental biology; cell signalling; insecticide;  
KW prokaryota; developmental biology; cell signalling; insecticide;  
XX pharmaceutical; gene; ds.  
XX

XX *Drosophila melanogaster*.  
PN WO200171042-A2.

EN MO2001/1012-AZ.  
 XX  
 XX  
 PD 27-SEP-2001.

PD	27-SEP-2001.
XX	
XX	
PF	23-MAR-2001: 2001WO-US009231.

23-MAR-2001; 2001WO-US009231.  
PF  
XX  
23 MAR 2000 2000WZ 010152Z  
DN

XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
XX  
PA (PEKE ) PE CORP NY.

XX  
PI  
XX

Venter JC, Adams M, Li PWD, Myers EW;

XX	WPI; 2001-656860/75.
XX	
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more
PT	genes from Drosophila and for elucidating cell signalling and cell-cell
PT	interactions.
XX	
XX	Claim 1; SEQ ID NO 5728; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent
CC	capable of detecting 1000 or more genes from Drosophila. The invention is
CC	useful in developmental biology and in elucidating cell signalling and
CC	cell-cell interactions in higher eukaryotes for the development of
CC	insecticides, therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC	sequences (ABL01840-ABL16175) and the encoded proteins (ABR57737-
CC	ABR72072). The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from <a href="http://wipo.int/pub/published_pct_sequences">wipo.int/pub/published_pct_sequences</a>
XX	
SQ	Sequence 2109 BP; 485 A; 589 C; 605 G; 430 T; 0 U; 0 Other;

Query Match	5.6%	Score 136.2	DB 4	Length 2109
Best Local Similarity	46.4%	Pred. No. 1.1e 19		
Matches 607	Conservative	0	Mismatches 678	Indels 24
				Gaps 4

Matches	607	Conservative	0	Mismatches	678	Indels	24	Gaps	4
Qy	668	TGTTTGCACTCTTGCCCTATAACTGGAAC	TGCTGGTTTATACACTGCGCCTCGTCTTCC	727					

Qy 668 TGCTTGTCAC TCTTGCC TATAACTGGA ACTGCTGGTTTATACCACTGCGCC TCGTCTTCC 727

[illegible]

Db	1076	AGACGGTGATGGCCTCCCTCGAAGCGTGGTTCTCTATTTTGGCGAGATTAGTATACTCAATA	1133
Qy	1793	CAGG---AGGAGGAAACCGTCGAACTGCCAAATGTGTGGTGGCCCAACGGGTTTGGCCAATCTTT	1849
Db	1136	TGGGCACCGCAGGCAACCGACGACAGCCAGCGTTCGCTCAGTGGGATACAGCGCCCTCT	1195
Qy	1850	TAACTCTAGACAAAAGACCCCTCCAAGAAATTCAGTGCAATATCCAGATTCTGCAAGGA	1909
Db	1196	TCGTCTCTGACGAAGAAGGACATGTGGGACGTGCTGAAGAGATATCCGGCGCCGCTGTTTC	1255
Qy	1910	TCCTCATGAAAGACGACGAGTGCTTTAAAGCAGAAGGCTTAAGACCGC	1958
Db	1256	GTCTGGAGTCGATAGCGTCAAGCGATTGAGAGAAATACAAGAAAGGCCCC	1304
RESULT 15			
ADFS6505			
ID ADFS6505 standard; DNA; 3027 BP.			
XX	ADFS6505;		
AC			
XX			
XX	12-FEB-2004 (first entry)		
XX	Modified rat olfactory cyclic nucleotide-gated ion channel DNA #1.		
DE			
XX	ds; gene; olfactory cyclic nucleotide-gated ion channel;		
KW	increased cAMP sensitivity; decreased cGMP sensitivity;		
KW	decreased nitric oxide sensitivity;		
KW	decreased calcium-calmodulin sensitivity.		
XX			
OS	Rattus norvegicus.		
XX			
XX	Key Location/Qualifiers		
PH	340..2334		
FT	CDS		
FT	/*tag= a		
XX			
XX	US2003157571-A1.		
XX			
XX	21-AUG-2003.		
XX			
PF	15-NOV-2002; 2002US-00295573.		
XX			
PR	16-NOV-2001; 2001US-0332494P.		
XX			
PA	(COLS ) UNIV COLORADO.		
XX			
FI	Karpen JW, Rich TC, Cooper DMF, Schaack J;		
XX			
XX	WPI; 2003-787336/74.		
DR	P-PSDB; ADFS6501.		
XX			
PT	New nucleic acid encodes modified olfactory cyclic nucleotide-gated ion		
PT	channels which exhibit increased sensitivity and specificity for cAMP and		
PT	signaling.		
XX			
PS	Disclosure; SEQ ID NO 1; 63pp; English.		
XX			
CC	The invention relates to an isolated nucleic acid encoding a modified		
CC	olfactory cyclic nucleotide-gated ion channel, where the channel		
CC	comprises mutations which together impart increased cAMP sensitivity,		
CC	decreased cGMP sensitivity, decreased nitric oxide sensitivity and		
CC	decreased calcium-calmodulin sensitivity. The invention is used to		
CC	elucidate activities of receptors, G-proteins, phosphodiesterases, adeny		
CC	cyclases and other proteins important in cAMP signaling. The present		
CC	sequence represents modified rat olfactory cyclic nucleotide-gated ion		
CC	channel DNA.		
XX			
SQ	Sequence 3027 BP; 783 A; 698 C; 746 G; 800 T; 0 U; 0 Other;		

```
Query Match      5.5%; Score 133.2; DB 10; Length 3027;
Best Local Similarity 46.8%; Pred. No. 5.7e-19;
Matches 623; Conservative 0; Mismatches 673; Indels 36;
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2005, 16:53:01 ; Search time 8615 Seconds  
(without alignments)  
13197.053 Million cell updates/sec

Title: US-09-855-828-3  
Perfect score: 2430  
Sequence: 1 atgtttaatcgtgacaaa.....aagaaaggctaaagcaataa 2430

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1707.4	70.3	2103	11	DQ043744 Homo sapi
2	1552.6	63.9	2103	11	DQ043745 Pan trogl
3	815.8	33.6	2056	4	AK040140 Mus muscu
4	553	22.8	553	2	BF725470 bxl6c03.y
5	478	19.7	527	1	AL113036 DKEP686A
6	477.4	19.6	1947	4	CR626336 full-leng
7	470	19.3	655	2	BB635459 BB635459
8	457	18.8	3444	11	DQ032773 Homo sapi
9	430.2	17.7	437	2	BF725469 bxl6c03.x
10	402.4	16.6	881	2	BE287002 601097229
11	394.6	16.2	1455	8	DN691557 CGX77-A10
12	388.8	16.0	1377	8	DN679703 CGX77-F03
13	373.4	15.4	1140	8	DN686265 CGX46-A07
14	363.8	15.0	1054	5	BX401127 BX401127
15	333.8	13.7	3444	11	DQ032774 Pan trogl
16	322	13.3	1356	8	DN680109 CGX10-A08
17	320.8	13.2	468	1	AA069559 zf75a04.x
18	320.8	13.2	656	3	BP224184 BP224184
19	308.4	12.7	737	8	DN900758 naql0c08.
20	308.4	12.7	978	3	BQ068992 AGENCOURT
21	304	12.5	633	7	CK619966 mk30e03.y
22	301.4	12.4	694	1	AU137778 AU137778

ALIGNMENTS

RESULT 1	DQ043744	2103 bp	DNA	linear	GSS 02-JUN-2005
LOCUS	DQ043744	Homo sapiens CNGB3 gene, VIRTUAL TRANSCRIPT, partial sequence.			
DEFINITION	DQ043744	Homo sapiens CNGB3 gene, VIRTUAL TRANSCRIPT, partial sequence.			
ACCESSION	DQ043744	GI:66894959			
VERSION	DQ043744	GSS			
KEYWORDS	DQ043744	Homo sapiens (human)			
SOURCE	DQ043744	Homo sapiens			
ORGANISM	DQ043744	Homo sapiens			
REFERENCE	DQ043744	1 (bases 1 to 2103)			
AUTHORS	DQ043744	Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B., Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.			
TITLE	DQ043744	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees			
JOURNAL	DQ043744	(er) PLoS Biol. 3 (6), E170 (2005)			
PUBMED	DQ043744	15869325			
REFERENCE	DQ043744	2 (bases 1 to 2103)			
AUTHORS	DQ043744	Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B., Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civello,D., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.			
TITLE	DQ043744	Direct Submission			
JOURNAL	DQ043744	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA			
COMMENT	DQ043744	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.			
FEATURES	DQ043744	Location/Qualifiers			
source	DQ043744	1. 2103			
organism	DQ043744	/organism="Homo sapiens"			
mol_type	DQ043744	/mol_type="genomic DNA"			
db_xref	DQ043744	/db_xref="taxon:9606"			
chromosome	DQ043744	/chromosome="8"			
gene	DQ043744	<1..2103			
locus_tag	DQ043744	/locus_tag="HC10349"			
ORIGIN	DQ043744				
Query Match	DQ043744	70.3%	Score 1707.4;	DB 11;	Length 2103;
Best Local Similarity	DQ043744	81.2%	Pred. No. 0;		
Matches 1708;	DQ043744	Conservative	0;	Mismatches 395;	Indels 0;
Gaps	DQ043744				0;

QY 1 ATGTTTAAATCGCTGACAAAGTCAACAAAGGTGAAGCTTATAGGAGAGAAACAATGAGAT 60  
DB |||||  
QY 1 ATGTTTAAATCGCTGACAAAGTCAACAAAGGTGAAGCTTATAGGAGAGAAACAATGAGAT 60  
DB |||||  
QY 61 GAACAAGTTCTCGTCGGATGAAGAAGCTCTCACCAGTAATCAGTCTCAGCAACC 120  
DB |||||  
QY 61 GAACAAGTTCTCGTCGGATGAAGAAGCTCTCACCAGTAATCAGTCTCAGCAACC 120  
DB |||||  
QY 121 ACAGCAG 180  
DB |||||  
QY 121 ACAGCAG 180  
DB |||||  
QY 181 ACGTCTGAAGCCACACACCAACATACAAAGACAACTCTCCAAGAGAAATTCCTCTGA 240  
DB |||||  
QY 181 NNN 240  
DB |||||  
QY 241 GATCTGACCACAAACCTGACCTCAAAATGAGCGAGAACCAACTGGAAACAGTGCCAGAG 300  
DB |||||  
QY 241 GATCTGACCACAAACCTGACCTCAAAATGAGCGAGAACCAACTGGAAACAGTGCCAGAG 300  
DB |||||  
QY 301 CAGAAAGAAATGGACCCCGGAGAGAGGTCCAAACAGCGCCCAAAAACAAACCGCTGCA 360  
DB |||||  
QY 301 CAGAAAGAAATGGACCCCGGAGAGAGGTCCAAACAGCGCCCAAAAACAAACCGCTGCA 360  
DB |||||  
QY 361 GCTCCTGTTTAAATGAGTATGCCGATGCCAGTCCAGCTACACAACTGGTGAAGAGAAATGCGT 420  
DB |||||  
QY 361 GCTCCTGTTTAAATGAGTATGCCGATGCCAGTCCAGCTACACAACTGGTGAAGAGAAATGCGT 420  
DB |||||  
QY 421 CAAAGAACAGCCCTCTACAGAGAAAAGTTGGTAGAGGGAGATCTCTCTCACCCGAGCC 480  
DB |||||  
QY 421 CAAAGAACAGCCCTCTACAGAGAAAAGTTGGTAGAGGGAGATCTCTCTCACCCGAGCC 480  
DB |||||  
QY 481 AGCCACACAACTGCAAGAGCCACCGCTGTACACAGTAAAGAGAGCGATCATAGCCA 540  
DB |||||  
QY 481 AGCCACACAACTGCAAGAGCCACCGCTGTACACAGTAAAGAGAGCGATCATAGCCA 540  
DB |||||  
QY 541 ACAGAACATTAACAGAGCTGTGTGGTTCAAAGTCAAAAGATGCTTTAAACAGAGTAC 600  
DB |||||  
QY 541 ACAGAACATTAACAGAGCTGTGTGGTTCAAAGTCAAAAGATGCTTTAAACAGAGTAC 600  
DB |||||  
QY 601 TTAAGCGAATTAACCTTCCAAACAGCATAGATTATACACAGATCGACTATCTCCTG 660  
DB |||||  
QY 601 TTAAGCGAATTAACCTTCCAAACAGCATAGATTATACACAGATCGACTATCTCCTG 660  
DB |||||  
QY 661 TGGCTCTTGCTGTCACTCTTGCTTAACTGGAAGTGTGGTTTATACCACTGGCCCTC 720  
DB |||||  
QY 661 TGGCTCTTGCTGTCACTCTTGCTTAACTGGAAGTGTGGTTTATACCACTGGCCCTC 720  
DB |||||  
QY 721 GTCCTCCCATATCAAAACGAGACAACTACTACTGCTTTATTCGGACATCATATGT 780  
DB |||||  
QY 721 GTCCTCCCATATCAAAACGAGACAACTACTACTGCTTTATTCGGACATCATATGT 780  
DB |||||  
QY 781 GATATCATCTACTTTATGATATGCTATTTATCCAGCCAGACTCGAGTTTGTAGAGGA 840  
DB |||||  
QY 781 GATATCATCTACTTTATGATATGCTATTTATCCAGCCAGACTCGAGTTTGTAGAGGA 840  
DB |||||  
QY 841 GGAGACATAATAGTGGATTCAATGAGCTTAAGGAAACACTACAGGACTTCTACAAAATTT 900  
DB |||||  
QY 841 GGAGACATAATAGTGGATTCAATGAGCTTAAGGAAACACTACAGGACTTCTACAAAATTT 900  
DB |||||  
QY 901 CAGTTGGATGTCGCATCAATAATACCAATTTGATATTTGCTACTCTCTCTTTGGGTTTAA 960  
DB |||||  
QY 901 CAGTTGGATGTCGCATCAATAATACCAATTTGATATTTGCTACTCTCTCTTTGGGTTTAA 960  
DB |||||  
QY 961 CCAATGTTTAGAGCAATAGGATGTTAAAGTACATCTCATTTTTTGAATTTAATCATCAC 1020  
DB |||||  
QY 961 CCAATGTTTAGAGCAATAGGATGTTAAAGTACATCTCATTTTTTGAATTTAATCATCAC 1020  
DB |||||  
QY 1021 CTAGAGTCTATAATGGACAAAGCATATATCTACAGAGTTATTTCGAACAACTGGATCTG 1080  
DB |||||  
QY 1021 NNN 1080  
DB |||||

QY 1081 CTGTTTATTCTGCACATTAATGCTCTGCTTTATTACTCGGCTTCAAACTATGAAGGAAAT 1140  
DB NNN 1140  
QY 1141 GGCACCTACTAGATGGGTGTATGATGGGAGAGAAACGAGTATCTGAGATGTTTATTATGG 1200  
DB |||||  
QY 1141 NNN 1200  
DB |||||  
QY 1201 GCAGTTCGAACCTTTAAATACCATTTGGTGGCTTCCAGAACCAACAACTTTATTTGAAAT 1260  
DB |||||  
QY 1201 GCAGTTCGAACCTTTAAATACCATTTGGTGGCTTCCAGAACCAACAACTTTATTTGAAAT 1260  
DB |||||  
QY 1261 GTTTTTCAACTCTTGAAATTTTTTTCTGGAGTTTTTGTCTCCAGTTTAAATTCGTGAG 1320  
DB |||||  
QY 1261 GTTTTTCAACTCTTGAAATTTTTTTCTGGAGTTTTTGTCTCCAGTTTAAATTCGTGAG 1320  
DB |||||  
QY 1321 ATGAGAGATGATTTGGAGCAGCTACAGCCAACTAGAACTACTTCCGGCCCTGCATGAT 1380  
DB |||||  
QY 1321 ATGAGAGATGATTTGGAGCAGCTACAGCCAACTAGAACTACTTCCGGCCCTGCATGAT 1380  
DB |||||  
QY 1381 GACACATTTGCTACATGAACAAATTTACTTCTTAACTTGTGCAAAAGCGAGTTCCG 1440  
DB |||||  
QY 1381 GACACATTTGCTACATGAACAAATTTACTTCTTAACTTGTGCAAAAGCGAGTTCCG 1440  
DB |||||  
QY 1441 ACTTGGTATGAATATACATGGGACTCTCAAAGAACTAGNNNNNNNNNNNNNNNNNN 1500  
DB |||||  
QY 1441 ACTTGGTATGAATATACATGGGACTCTCAAAGAACTAGNNNNNNNNNNNNNNNNNN 1500  
DB |||||  
QY 1501 ACCCTACCACTACGGTCCAGTTAGCCCTCGCATTTGATGTGAACCTTCAGCATCATCAG 1560  
DB |||||  
QY 1501 NNN 1560  
DB |||||  
QY 1561 AAAGTCGACTTGTTCAGGGTTGTGATACACAGATGATTTATGACATGTTGCTAAGATTG 1620  
DB |||||  
QY 1561 NNN 1620  
DB |||||  
QY 1621 AAATCCGTTCTCTATTTGCTTGGTGAATTTGCTGCAAAAAGGGAGAAATTTGGCAAGAA 1680  
DB |||||  
QY 1621 AAATCCGTTCTCTATTTGCTTGGTGAATTTGCTGCAAAAAGGGAGAAATTTGGCAAGAA 1680  
DB |||||  
QY 1681 ATGTATATCATCAAGCATGGAGAACTCAAGTTCTTGGAGCCCTGATGCTACTAAAGTT 1740  
DB |||||  
QY 1681 ATGTATATCATCAAGCATGGAGAACTCAAGTTCTTGGAGCCCTGATGCTACTAAAGTT 1740  
DB |||||  
QY 1741 CTGGTTACTCTGAAAGCTGGGTGGTGTTCGAGAAATCAGCTTCTTAGCTAGCTAGCTA 1800  
DB |||||  
QY 1741 CTGGTTACTCTGAAAGCTGGGTGGTGTTCGAGAAATCAGCTTCTTAGCTAGCTAGCTA 1800  
DB |||||  
QY 1801 GGAACCCGTGAACTGCCAATGTGGTGGCCCAACGGGTTTGGCAATCTTTTAACTCTAGAC 1860  
DB |||||  
QY 1801 GGAACCCGTGAACTGCCAATGTGGTGGCCCAACGGGTTTGGCAATCTTTTAACTCTAGAC 1860  
DB |||||  
QY 1861 AAAAGAGCCCTCCAAGAAATTTCTAGTGCATTTATCCAGATTTCTGAAGGATCCTCATGAAG 1920  
DB |||||  
QY 1861 AAAAGAGCCCTCCAAGAAATTTCTAGTGCATTTATCCAGATTTCTGAAGGATCCTCATGAAG 1920  
DB |||||  
QY 1921 AAAGCCAGAGTGTCTTTAAAGCAGAGGCTAAGACCGCAGAGCAACCCCTCCAAGAAAA 1980  
DB |||||  
QY 1921 AAAGCCAGAGTGTCTTTAAAGCAGAGGCTAAGACCGCAGAGCAACCCCTCCAAGAAAA 1980  
DB |||||  
QY 1981 GATCTTGCCTCTCTTCCCAACCGAAAGAGAGACACCCAACTGTGTAAAACTCTCCTA 2040  
DB |||||  
QY 1981 GATCTTGCCTCTCTTCCCAACCGAAAGAGAGACACCCAACTGTGTAAAACTCTCCTA 2040  
DB |||||  
QY 2041 GGAGCCACAGGAAAGAGAGTCTTGCAGACTACTCAAAATTTGAAGCGAGAGCAAGCAGCT 2100  
DB |||||  
QY 2041 GGAGCCACAGGAAAGAGAGTCTTGCAGAGTCTTGCAGAGTCTCAAAATTTGAAGCGAGAGCAGCT 2100  
DB |||||  
QY 2101 CAG 2103  
DB |||||  
QY 2101 CAG 2103  
DB |||||

```
RESULT 2
DQ043745
LOCUS
DEFINITION
  DQ043745 2103 bp DNA linear GSS 02-JUN-2005
  Pan troglodytes CNGB3 gene, VIRTUAL TRANSCRIPT, partial sequence,
  genomic survey sequence.
ACCESSION
  DQ043745
VERSION
  DQ043745.1 GI:66894960
KEYWORDS
  GSS.
SOURCE
  Pan troglodytes (chimpanzee)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
  Homnidae; Pan.
REFERENCE
  1 (bases 1 to 2103)
  Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
  Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Cividello, D.,
  White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
  A Scan for Positively Selected Genes in the Genomes of Humans and
  Chimpanzees
  (er) PLOS Biol. 3 (6), E170 (2005)
  15869325
REFERENCE
  2 (bases 1 to 2103)
  Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B.,
  Hubisz, M.J., Fiedel-Alon, A., Tanenbaum, D.M., Cividello, D.,
  White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
  Direct Submission
  Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
  Rockville, MD 20850, USA
COMMENT
  This sequence was made by sequencing genomic exons and ordering
  them based on alignment. Translation starts at the beginning of
  alignment.
FEATURES
  source
    1. .2103
    /organism="Pan troglodytes"
    /mol_type="genomic DNA"
    /db_xref="taxon:9598"
    <1..>2103
    /gene="CNGB3"
    /locus_tag="HC10349"
ORIGIN
  Query Match 53.98; Score 1552.6; DB 11; Length 2103;
  Best local Similarity 75.44; Pred. No. 0;
  Matches 1587; Conservative 0; Mismatches 513; Indels 6; Gaps 2;
  QY 1 ATGTTTAAATCGCTGACAAAGTCAACAAGTGAAGCTATAGAGAGAAACAATGAGAAT 60
  DB 1 ATGTTTAAATCGCTGACAAAGTCAACAAGTGAAGCTATAGAGAGAAACAATGAGAAT 60
  QY 61 GAACAAAGTTCTCGTGGAAATGAAGAGGCTCTCACCAAGTAATCACTCAGCAAAACC 120
  DB 61 GAACAAAGTTCTCGTGGAAATGAAGAGGCTCTCACCAAGTAATCACTCAGCAAAACC 117
  QY 121 ACAGCAGAG--GAAGAAACAAGGTGAAGAGAAATCTCTCAAAACCAAGTCAATCCA 177
  DB 118 ACAGCAGAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 177
  QY 178 GTCACGCTGAGAGGCGACACCAACATACAGACAAACTCTCAAGAAATAATTCCTCT 237
  DB 178 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 237
  QY 238 GGAGATCTGACCAAAACCCCTGACCCCTCAAAATGACAGCAAGCACTGGAACAGTGCCA 297
  DB 238 GGAGATCTGACCAAAACCCCTGACCCCTCAAAATGACAGCAAGCACTGGAACAGTGCCA 297
  QY 298 GAGCAGAGGAATGGAGCCCGGAAAGAGGTCCAAACAGGCCCAACAAACCAACCGCCT 357
  DB 298 GAGCAGAGGAATGGAGCCCGGAAAGAGGTCCAAACAGGCCCAACAAACCAACCGCCT 357
  QY 358 GCAGCTCTGTTAATAATGAGTATGCCGATGCCGAGCTACACACCTGGTGAAGAAGATG 417
  DB 358 GCAGCTCTGTTAATAATGAGTATGCCGATGCCGAGCTACACACCTGGTGAAGAAGATG 417
  QY 418 CGTCAAGAAACAGCCCTCTACAAGAAAAAGTTGGTAGGGAGATCTCTCTCACC CGAA 477
  DB 418 CGTCAAGAAACAGCCCTCTACAAGAAAAAGTTGGTAGGGAGATCTCTCTCACC CGAA 477
  QY 478 GCCAGCCCAACAACTGCAAGAGCCCGCTGTACCAAGTAAAGAGGAGCGATGATAAG 537
  DB 478 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 537
  QY 538 CCAACAGAACATTACTACAGGCTGTGTGGTTCAAAGTCAAAAGATGCTTTTAACAGAG 597
  DB 538 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 597
  QY 598 TACTTAAAGCGAATTAAGCTTCCAAACAGATAGATTATACACAGATCGACTCTATCTC 657
  DB 598 TACTTAAAGCGAATTAAGCTTCCAAACAGATAGATTATACACAGATCGACTCTATCTC 657
  QY 658 CTGTGGCTCTTCTGTCTCACTCTTGCCTATAACTGGAACCTGCTGTTTATACCACTGC 717
  DB 658 CTGTGGCTCTTCTGTCTCACTCTTGCCTATAACTGGAACCTGCTGTTTATACCACTGC 717
  QY 718 CTGCTCTTCCCATATCAAAACCGAGACAACTACTACTGCTTATTCGGGACATCATAT 777
  DB 718 CTGCTCTTCCCATATCAAAACCGAGACAACTACTACTGCTTATTCGGGACATCATAT 777
  QY 778 TGTGATATCATCTACCTTTATGATATGCTATTTATCCAGCCAGACTCCAGTTTGAAGA 837
  DB 778 TGTGATATCATCTACCTTTATGATATGCTATTTATCCAGCCAGACTCCAGTTTGAAGA 837
  QY 838 GGAGGAGACATAATAGTGGATTCAATGAGCTAAAGGAAACACTACAGGACTTCTACAAA 897
  DB 838 GGAGGAGACATAATAGTGGATTCAATGAGCTAAAGGAGACTACAGGACTTCTACAAA 897
  QY 898 TTTCAAGTTGGATGTGCGATCAATAATACATTTGATATTTGCTACCTCTCTTTGGGTTT 957
  DB 898 TTTCAAGTTGGATGTGCGATCAATAATACATTTGATATTTGCTACCTCTCTTTGGGTTT 957
  QY 958 AATCCAAATTTAGAGCAAAATAGGATGTTAAAGNNNNNNNNNNNNNNNNNNNNNNNN 1017
  DB 958 AATCCAAATTTAGAGCAAAATAGGATGTTAAAGNNNNNNNNNNNNNNNNNNNNNNNN 1017
  QY 1018 CACCTAGAGTCTATAATGGACAAAGCATATATCTACAGAGTTATTCCGAAACACTGGATC 1077
  DB 1018 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1077
  QY 1078 TTGCTGTTTATCTGACATTAATGCTGTGTTTATTACTGGGCTTCAAACATGAGGA 1137
  DB 1078 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1137
  QY 1138 ATTGGCACTACTAGATGGGTGATGATGGGAGGAAACGAGTATCTGAGATCTTATTAT 1197
  DB 1138 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1197
  QY 1198 TGGGCGAGTTTCAACTTTAAATACCATTTGGTGGCTTCCAGAACCAAACTTTATTGAA 1257
  DB 1198 TGGGCGAGTTTCAACTTTAAATACCATTTGGTGGCTTCCAGAACCAAACTTTATTGAA 1257
  QY 1258 ATTTGTTTTCAACTCTTGAATTTTTCTGGAGTTTCTGTTCTCCAGTTTAAATGGT 1317
  DB 1258 ATTTGTTTTCAACTCTTGAATTTTTCTGGAGTTTCTGTTCTCCAGTTTAAATGGT 1317
  QY 1318 CAGATGAGAGATGTTGATGGAGCAGCTACAGCCCAATCAGAACTACTCCGCGCTGCATG 1377
  DB 1318 CAGATGAGAGATGTTGATGGAGCAGCTACAGCCCAATCAGAACTACTCCGCGCTGCATG 1377
  QY 1378 GATGACACCAATTCCTTACATGAACAATTAATCTCAATCTTCTAACTTGTGCAAAAGCGAGTT 1437
  DB 1378 GATGACACCAATTCCTTACATGAACAATTAATCTCAATCTTCTAACTTGTGCAAAAGCGAGTT 1437
  QY 1438 CGGACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTAGATGAGTCTGATTTGCTT 1497
  DB 1438 CGGACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTAGNNNNNNNNNNNNNN 1497
  QY 1498 AAGACCCCTACCAACTAGCGTCCAGTTAGCCCTCGCCATTTGATGTGAACACTTCAGCATCATC 1557
```



```
QY 1136 GAATTGGCACTACTAGATGGGTGATGATGGGGAAGAAACGAGTATCTGAGATGTTATT 1195
Db      |||||
QY 62  GAATTGGCTCAACTAAATGGGTCTATAATGGTGAAGCAACAAGTATCTGCGATGCTTTT 121
Db      |||||
QY 1196 ATTGGGCACTCGAAGCTTTAATTACCATTTGGTGGCTTCCAGAACACAAACTTTATTG 1255
Db      |||||
QY 122  ATTGGGCACTCGAAGCTTTAATTACTATCGGGGCTTCCAGAGCCACAGACTTCATTG 181
Db      |||||
QY 1256 AAATTGTTTTTCAACTCTTGAATTTTTTTTCTGGAGTTTTTTGTGTTCTCCAGTTAAATTG 1315
Db      |||||
QY 182  AAATTGTTTTTCAACTCTTGAATTTTTTTTCTGGGGTTTTTGTGTTCTCCAGCTTAATTG 241
Db      |||||
QY 1316 GTCAGATGAGAGATGATGAGAGAGCTACAGCCAAATCAGAACTACTTCCGGGCTGCA 1375
Db      |||||
QY 242  GTCAGATGGGTGATGATGATGGGCGACCAACAGCCAAATCAGAACTACTTCCAGGCTGCA 301
Db      |||||
QY 1376 TGGATGACACCACTTGCCTACATGAACAATTTACTTCCATTTCTTAACTTGTGCAAAACGAG 1435
Db      |||||
QY 302  TGGACCATATCATTTGCTTACATGAACAATACTCTATTTCTCAGAGTGTGAGTATCGAG 361
Db      |||||
QY 1436 TTCGGACTTGGTATGAATATACATGGGACTCTCAAGAATGCTAGAGTCTGTGATTTGC 1495
Db      |||||
QY 362  TTCGGACTTGGCTGGATATACATGGGACTCACAAGAAATCTTAGATGAGTCCAACTGC 421
Db      |||||
QY 1496 TTAAGACCCCTACCAACTACGGTCCAGTTAGCCCTCGCCATTTGATGTGAATTCAGCATCA 1555
Db      |||||
QY 422  TTGAGAACCTCCGACAGCAATGCAATGCTGTCTATTTGCCCTTGACATAAACTTCAGTATCA 481
Db      |||||
QY 1556 TCAGCAAGTCCGACTTTGTTCAAGGTTGTGATACAGATGATTTATGATGATGATGATGATGAT 1615
Db      |||||
QY 482  TCAGCAAGTGGAGTATTTCAAGGCTGTGACACAGATGATTTATGATGATGATGATGATGATGAT 541
Db      |||||
QY 1616 GATTGAAATCCGTTCTCTATTTTCCCTGGTGGTCTGTTCTCAAAAAGGAGAAATTTGGCA 1675
Db      |||||
QY 542  GATTGAAATCCACTATTTATTTTACCTGGTGGTCTGTTCTGCAAAAAGGAGAAATTTGGAA 601
Db      |||||
QY 1676 AGAAATGTTATATCATCAAGATGGAGAAATCCAAAGTTCTTGAGAGCCCTGATGTTACTA 1735
Db      |||||
QY 602  AAGAAATGTTATCATCATCAACAGGAGAAATCCAAAGTCTTGGAGGCCCTGATGTTGCTC 661
Db      |||||
QY 1736 AAGTTCTGTTTACTCTGAAAGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1795
Db      |||||
QY 662  AAGTTCTGTTTACTCTGAAAGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 721
Db      |||||
QY 1796 GAGGAGAAACCGTGGCAATGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1855
Db      |||||
QY 722  GAGGAGAAATCGCGGACAGCTGATGTTGTCGCCACGATTTGCCAATCTTTTAACTC 781
Db      |||||
QY 1856 TAGACAAAAGACCTCCAAAGAAATTTAGTGTGATTTCCAGATTTCTGAAAGGATCTTCA 1915
Db      |||||
QY 782  TGGACAAAAGACTCTCCAAAGAAATTTCTGCTGCAATTTCAAACTTTCAAAAAGCTCTCTCA 841
Db      |||||
QY 1916 TGAAGAAAGCAGAGTCTTTTAAAGCAGAGGCTAGACGGCAGAGCAACCTCTCAA 1975
Db      |||||
QY 842  TGAAGAAAGCAGAAATTTCTTTTAAAGCAGAGGGAAGCACTCAGGCAATCCCTGCAA 901
Db      |||||
QY 1976 GAAAGATCTTGGCCCTCTCTTCCACCGAAGAGAGACACCCAAACTGTTTAAAACTC 2035
Db      |||||
QY 902  GACGAGACCTGCTCTTTTCCACCAAAAGAGAGACACCCCGATGCTTAAAGTTC 961
Db      |||||
QY 2036 TCCTAGGAGGACAGGAAAGCAAGTCTTCCAGACTTCTCAAGTCTCAATTTGAAGCGAGCAAG 2095
Db      |||||
QY 962  TCCTAGGAAACACAGGAAAGGTGGACCTTGGAGACTCTTAAAGAAAGAGAGAAACAA 1021
Db      |||||
QY 2096 CAGCTCAGAGAAAGAAATTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2155
Db      |||||
QY 1022 CCACTCAGAA-----ATAAAGGTGAAA 1043
Db      |||||
QY 2156 AACAAAAAGAAATGAAGATATAACAAAAAGAAATGAAGTAAAGGAAAGAAATGAAG 2215
Db      |||||
QY 1044 GTTCCAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1103
Db      |||||
```

```
QY 2216 ATAAAGATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2275
Db      |||||
QY 1104 AAAAAGAGA-----AAAAAACCCACTGGATAGTTTCAAGAGCACTAA 1149
Db      |||||
QY 2276 GTCTTATTGCACTGGAGGAAGAACCCCACTAGTTAGAGGACAGTTTTTACCCAGAGGA 2335
Db      |||||
QY 1150 GTTCTATTACAGTGGAGGAATGTCCCAATCAATTAATAATAGGAGTTCCTCCCAAGGAC 1209
Db      |||||
QY 2336 CTTCTCGTCAATCACTATTATCAGATGGCTCTTCTGCTGAGGCGGAGAGAGAGTTTC 2395
Db      |||||
QY 1210 ATGCTTATCAATTTGATCATCAGCAATAGG-TCCTGCTACTGAGACTGGAGAGAGATTTC 1268
Db      |||||
QY 2396 TTACTATTGAAGTCAAGAAAAAGGCTTAAGCAATAA 2430
Db      |||||
QY 1269 TGACTATCAAGTCAAGAAAAAGGCTTAAGCAATAA 1303
Db      |||||
```

RESULT 4  
BF725470 553 bp mRNA linear EST 05-JAN-2001  
LOCUS bx16c03.y1 Human Iris cDNA (Un-normalized, unamplified): BX Homo  
DEFINITION sapiens cDNA clone bx16c03 5', mRNA sequence.  
ACCESSION BF725470  
VERSION BF725470.1 GI:12041381  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 553)  
Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.  
NEIBANK: EST analysis and bioinformatics for ocular genomics  
Invest. Ophthalmol. Vis. Sci. 41 (2000) in press  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 16 row: c column: 03  
Seq primer: M13RP1 reverse primer (ABI).  
FEATURES  
Location/Qualifiers  
1..553  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="bx16c03"  
/tissue\_type="Iris"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human Iris cDNA (Un-normalized, unamplified):  
BX"  
/note="Organ: Eye; Vector: pCMVSPORT6; Post-mortem iris  
tissue was pooled from 10 individuals ranging in age from  
4-80 years and RNA was extracted. From this pooled sample  
an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A  
directionally cloned cDNA library in the pCMVSPORT6 vector  
was constructed at Life Technologies, essentially  
following the protocols of the SuperScript Plasmid System  
full details of which are contained in the manufacturer's  
instruction manual (<http://www.lifetech.com/>). First  
strand synthesis was carried out using a Not I  
primer-adaptor  
[5'-PGACTAGTTCTAGATCGGAGCGGCCGCC(T)15-3']. Not I/blunt  
end inserts were cloned into the Not I/EcoR V sites in the  
vector. EST analysis was performed on the unamplified  
library at the NIH Intramural Sequencing Center (NISC)."

Query Match 22.8%; Score 553; DB 2; Length 553;  
Best Local Similarity 100.0%; Pred. No. 1.1e-117;  
ORIGIN



**AUTHORS**  
**TITLE**  
**JOURNAL**

**COMMENT**

**FEATURES**  
 source  
 1. .1947  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DK009YE13"  
 /issue type="HeLa cells Cot 25-normalized"  
 /plasmid="pCMVSPORT\_6"

**ORIGIN**

Query Match 19.6%; Score 477.4; DB 4; Length 1947;  
 Best Local Similarity 64.6%; Pred. No. 5.7e-100;  
 Matches 712; Conservative 0; Mismatches 391; Indels 0; Gaps 0;

QY 861 AAATGAGCTAAGGAACACATACAGGACTTCTACAAAATTTTCAGTTGGATGTCGCATCAAT 920  
 DB 4 AAAGGACATCGGAAATAAATACCTGAAGTCTCGCCGCTTCAAGATGGACCTGCTCAGCCT 63

QY 921 AATACCAATTTGATATTTGCTACCTCTCTTTGGGTTTAAATCCAAATGTTTAGAGCAAAATAG 980  
 DB 64 CTGCGCCCTGGATTTCTCTATTTGAAGTCGGTGTGAACCCCTCTCTCGCGCTGCCCG 123

QY 981 GATGTTAAAGTACACTTCAATTTTGAATTAATCATCACTAGAGTCTATTAATGACAA 1040  
 DB 124 CTGTTTAAAGTACATGCGCTTCTCGAGTTTAAACAGCGCTCGAATCCATCTCAGCAA 183

QY 1041 AGCATATATCTACAGAGTATTTCGAACAACCTGATATCTGCTGTTTATTTCTGCACATTA 1100  
 DB 184 AGCTACGTGTACAGGCTCATCAGGACCAAGCTACCTTCTCTACAGCGCTGCAATTTGAA 243

QY 1101 TGCTGTGTTTATTACTGGGCTTCAAACTATGAAGAAATGGCACTACTAGATGGGTGTA 1160  
 DB 244 TTCTGTCTTTTATTACTGGGCATCGGCCTTACGGCTCGGCTCCACTCACTGGGTTA 303

QY 1161 TGATGGGAGGAACACAGTATCTGAGATGTTATTTGGGCGAGTTGCAATTTTAATAC 1220  
 DB 304 CGATGGCGTGGGAAACAGTTATATTCGCTGTACTTCTTGTGTGAAGACCCCTCATCAC 363

QY 1221 CATTTGGGCTTCCGAAACCAAACTTTATTTGAAATTTGTTTCAACTCTTGAATTT 1280  
 DB 364 CATCGGGGGCTGCTGACCCCAAGACACTCTTTTGAATTTGCTTCCAGCTGCTGAATTA 423

QY 1281 TTTTCTGGAGTTTGTGTTCTCCAGTTTAAATTTGGTCAGATGAGATGTTGATGGAGC 1340  
 DB 424 TTTACGGGGCTCTTCTGCTTCTGTGATGATCGGACATGAGATGTTGTTAGGGC 483

QY 1341 AGCTACAGCAATCAGAACTACTTTCGCGCTGTCATGATGATGACACCAATTCCTACATGAA 1400  
 DB 484 CGCACCGCGGACAGACCTTACTACCGCAGCTGATGACAGCAGCGGTGAAGTACATGAA 543

QY 1401 CAATTACTCATTTCTTAACTTTGCAAAAGCGAGTTTCGAGCTTGGTATGAATATACATG 1460  
 DB 544 TTTCTACAAGATCCCCAAGTCCGTCGAGAACCGCGTCAAGACCTGGTACGAGTACACCTG 603

QY 1461 GGAATCTCAAGAAATGCTAGATGATGTTGTTTGAAGACCTTACCAACTCGGTCCA 1520  
 DB 604 GCATCCGAGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663

QY 1521 GTTAGCCCTCGCCATTTGATGTAATTTTCAAGTATGATGATGATGATGATGATGATGATGAT 1580  
 DB 664 GCTGGACCTCGCCATCGAGTGAATCAACATGTTAGCAAAAGTTCGCACTCTTTCAAGG 723

QY 1581 TTGTGATACAGATGATTTTATGACATGTTGCTGAAGATTTGAATCCGCTTCTATTATGGCC 1640

Db 724 CTGTGACCGGAGATGATCTTTGACATCTGAAGAGGCTTCGCTCTGTGTCTACTCGCC 783

QY 1641 TGGTCACTTTGTCTCCAAAAGGAGAAATTTGGCAAGAAATGTATATCATCAAGCATGG 1700

Db 784 CAACGACTATGTGTGCAAGAAGGGGAGATCGGCGTGTGAGATGTATCATCATCCAGCAG 843

QY 1701 AGAAGTCCAAAGTTCTTGGAGGCCCTGATGATTAAGTTTCTGTTTACTTCTCAAGCTGG 1760

Db 844 GCAAGTGCAGGTCTTGGGCGGCCCTGATGGGAAATCTGTCTGTGACCTGAAAGCTGG 903

QY 1761 GTCGGTGTGAGAAATCAGCCTTCTAGCAGCAGGAGGAGAAACCGTTCGAATTCCTCAA 1820

Db 904 ATCTGTGTTTGGAGAAATAAGCTTCTGCTGCTTGGGCGGGAACCGCGCACGCCCCAA 963

QY 1821 TGTGTGGCCCAAGGGTTTGGCAATCTTTTAACTCTAGACAAAGAACCTTCAAGAAAT 1880

Db 964 CGTGTGGCGCACGGGTTTACCACCTCTTCTCATCTCTGGAATAAGAGGACCTTGAATGAGAT 1023

QY 1881 TCTAGTGCATTTCCAGATTTCTGAAGGATCTCATGAAGAGCCAGAGTCTCTTTTAA 1940

Db 1024 TTTGTGCAATATCTCTGAGTCTCAGAAGTTTCTCCGAAAGAACCGCGCATGCTGAG 1083

QY 1941 GCAGAAGGCTAAGACCGCAGAAAG 1963

Db 1084 AAGCAACAATAAGCCAGGAGG 1106

**RESULT 7**  
 BB635459  
 LOCUS  
 DEFINITION  
 BB635459  
 VERSION  
 BB635459.1 GI:16471509  
 KEYWORDS  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridae; Muridae; Mus.  
 REFERENCE  
 1 (bases 1 to 655)  
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,  
 Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,  
 Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,  
 Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,  
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
 Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
 Unpublished (2001)  
 CONTACT: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22, Suenho-cho, Tsukuba-shi, Ibaraki, 305-3858, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 waki, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
 and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.



[illegible]

QY	1877	AAATTC	TAGTGC	ATTATCC	GAGATCT	CGAAGG	ATCCTCAT	GAGAAAGCC	CAGAGTCT	TTT	1936
DB	3173	AGATTT	TGGTGC	ATTATCT	CGAGTCT	CAGAAAGT	TACTCG	GAAAGCC	AGCGCG	ATGC	3232
QY	1937	TAAAGC	AGAAGG	CTTAAG	ACCCG	CAGAAG					1963
DB	3233	TGAGAAG	CAACAAT	TAAGCC	CAAGGAGG						3259

RESULT 9	BF725469	437 bp	linear	EST 05-JAN-2001
LOCUS	BF725469/c			
DEFINITION	bx16c03.x1 Human Iris cDNA (Un-normalized, unamplified): BX Homo sapiens cDNA clone bx16c03 3'. mRNA sequence.			

ACCESSION	BF725469	
VERSION	BF725469.1	GI:12041380
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	

NCBI SOURCE	ORGANISM	EST.
Homo sapiens (human)	Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		

REFERENCE  
1 (bases 1 to 437)  
AUTHORS Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.  
TITLE NETBANK: EST analysis and bioinformatics for ocular genomics  
JOURNAL Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
COMMENT Contact: Wistow G

CONTACT: WISLOW G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452

Fax: 301 496 0078  
 Email: [graeme@helix.nih.gov](mailto:graeme@helix.nih.gov)  
 Plate: 16 row: c column: 03  
 Seq primer: -21M13 forward primer (ABI).

FEATURES	Location/Qualifiers
source	1. .437

```

1.11.10
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="bx16c03"
/tissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMDH108"
/clone_lib="Human Iris cDNA (Un-normalized, unamplified):
Rx"

```

Bx".  
/note="Organ: Eye; Vector: pCMVSPORT6; Post-mortem iris tissue was pooled from 10 individuals ranging in age from 4-80 years and RNA was extracted. From this pooled sample an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the SuperScript plasmid System full details of which are contained in the manufacturer's Instruction manual (<http://www.lifetech.com/>). First strand synthesis was carried out using a Not 1 primer-adaptor

primer-adaptor  
[5'-pGACTAGTTCTAGATCGGCGCGCCCTT(15-3')]. Not I/blunt  
end inserts were cloned into the Not I/EcoR V sites in the  
vector. EST analysis was performed on the unamplified  
library at the NIG Intramural Sequencing Center (NISC).

ORIGIN	Query Match	Score	DB 2:	Length	437:
	17.7%	430.2			

Query Match 17.7%; Score 430.2; DB 2; Length 437;  
Best Local Similarity 99.3%; Pred.No. 4.3e-89;

QY 1759 GGGTCGGTGTGGAGAAATCAGCCCTTCTACGACGAGGAGAGAAACCGTCGAATGCC 1818  
Matches 432; Conservative 0; Mismatches 3; Indels 0; Gaps 0

Db	437	GGGTCGGTGTTTGGAGAAATCAGCCTTC	TAGCAGCAGGAGGAGAACCGTCGAATGCC	378
Qv	1819	AATGTGGTGCCCGCCACGGGTTTGCCAA	TCTTTAACTCTAGACAAAMGACCCCTCCAMGA	1878



**AUTHORS**  
 Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,  
 Schmutz,J. and Myers,R.M.  
**TITLE**  
 Expressed sequence tags from *Gasterosteus aculeatus*  
**JOURNAL**  
 Unpublished (2003)  
**COMMENT**  
 Contact: Grimwood, Jane  
 Stanford Human Genome Center  
 Stanford University School of Medicine  
 975 S California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 320 5917  
 Fax: 650 320 5801  
 Email: jane@shgc.stanford.edu  
 Plate: 77  
 High quality sequence start: 18  
 High quality sequence stop: 941.  
**FEATURES**  
 source  
 1. 1455  
 /organism="Gasterosteus aculeatus"  
 /mol\_type="mRNA"  
 /strain="Comer Creek sticklebacks, WA USA"  
 /db\_xref="taxon:69293"  
 /clone="CGX77-A10"  
 /sex="mixed male and female"  
 /tissue\_type="eyes"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="SHGC-CGX"  
 /notes="Vector: Express 1; Total and poly A+ RNA was isolated from the indicated stickleback tissue, and a cDNA library was constructed in the Express 1 plasmid vector by Open Biosystems. First strand cDNA synthesis was primed with an 54 bp linker primer containing an oligodT sequence preceded by a synthetic NotI site (first strand primer: 5'-GACTAGTTCAGATCGAGCGGCCCTT)25-3'). Following second strand synthesis, cDNAs were made blunt at the end corresponding to the original 5 prime end of mRNA, and cloned directionally into the NotI and EcoRV sites of Express 1. Note that the EcoRV site is typically destroyed in the blunt end cloning, leaving a junction of the form 'xxxATC' (where is ATC is the second half of the EcoRV site, and xxx is derived from the cDNA sequence). A map of the Express 1 vector is available at: [http://www.openbiosystems.com/cdna\\_library\\_construction\\_fa\\_q.php#8](http://www.openbiosystems.com/cdna_library_construction_fa_q.php#8) The primary library was transformed and amplified in DH10B (T1 phage resistant) bacteria. Clones available from Open Biosystems: <http://www.openbiosystems.com/stickleback>"

**ORIGIN**  
 Query Match 16.2%; Score 394.6; DB 8; Length 1455;  
 Best Local Similarity 61.9%; Pred. No. 1.1e-80;  
 Matches 641; Conservative 0; Mismatches 389; Indels 6; Gaps 1;  
 QY 809 TTATCCAGCCAGACTCCAGTTTGTAGAGGAGGACATAATAGTGCATTCAAATGAGC 868  
 DB 39 TTGGCAGCCCGGCTACAGTTTGTCTAGCCGAGACGTTATTAAAGCGACCTATGG 98  
 QY 869 TAAGGAACACTACAGGACTTCTACAAATTTTCAGTTGGATGTCGATCAATATACCAT 928  
 DB 99 CTAAAGTATTATTACGGAAATCCCAACGATTAGATTGATATCTCAGCGTCTCGCCT 158  
 QY 929 TTGATATTGCTACTCTCTTTGGGTTTAATCCATGTTTATAGCAAAATAGGATGTAA 988  
 DB 159 TTGACCTTCTCGGCTGTACTTTGAGTTCTCCGCGGTATACAGAGTCAACCGCTTTATCA 218  
 QY 989 AGTACACTCATTTTGAATTTAATCATCATCCTAGAGTCTATATATGACAAACATATA 1048  
 DB 219 GGATTGAGTCTCTTTTGACTTTTATGATGATCGAGAGCATATGCGCAAGCCTACA 278  
 QY 1049 TCTACAGAGTTATTTCGAACAACTGGATCTTGCTGTTTATTTCGCACATTAAATCCCTGTG 1108  
 DB 279 TCTGGAGAGTGATTTCGACCAACAGGTTACCTGCTCTTCGCCCTCCATCTCACAGCTGTG 338  
 QY 1109 TTTTATTACTGGGCTTCAAACTATGAAGAAATGCGACTACTAGATGGGTGTATGATGGG 1168

DB 339 CGTACTAGTGGCTCCATGTACACGGGTATTAAGTCGACCACCGTGGGTGTATATGGAG 398  
 QY 1169 AAGGAAACAGTATCTGAGATGTTATTATTTGGGAGTTTGAACATTTTAATACCATTTGGTG 1228  
 DB 399 AGGGCAGCGGTACCTGCGTTGTTTACTACTACCGAGTCCGCACTGTGATCAACATCGGTG 458  
 QY 1229 GCCTTCCAGAACCAAACTTTATTGAAATTTGTTTTTCAACTCTTTGAATTTTTTTCTG 1288  
 DB 459 GTCTTCCAGAGCCAAATACCACTTTCGAGATATCTTTTCAGATGCGCAACTTTTTCATTG 518  
 QY 1289 GAGTTTTTGTCTCTCCAGTTTAAATTTGTCAGATCAGAGATGTCGATTTGGAGCAGCTACAG 1348  
 DB 519 GCGTTTTTGTATTCTCCAGTTTGAATTTGACAGATGTCATAGAGCGGCAACAG 578  
 QY 1349 CCAATCAGAACTACTTCCGCGCTGCATGGATGACACCATTCCTTACATGAAACAATTACT 1408  
 DB 579 CGGGGAAGACCTATTTCAGGGCATCCATGGTGGCTGTGTTGACTACATGAATACCTACA 638  
 QY 1409 CCATTCTTAACCTTGTGCNAAGCGAGTTCCGACTTGGTATGAATATACATGGGACTCTC 1468  
 DB 639 ATATCCCAAGGTTGTTTCAGAACAGAAATCCGGACCTGGGTACAACTACACCTGTCAGCTC 698  
 QY 1469 AAAGAATGCTAGATGAGTCTGATTTGCTTAAAGACCTACCAACTACGGTCCAGTTAGCCC 1528  
 DB 699 AGGGCATGCTGGACGAGTCGGAGCTGTAGACAGATGCCTCTGGTGATCAAACTGCCA 758  
 QY 1529 TCGCCATTGATGTGAATTTACAGCATCATCAGCAAAAGTCGACTTGTTCAGAGGTTGTGATA 1588  
 DB 759 TCGCGTGGACATCAACTTGTCAACTTCCAGAAGATTCATTGTTTCAGGGCTGTGACC 818  
 QY 1589 CACAGATGATTTATGACATGTTGCTTAAGATTGAAATCCGTTCTCTATTTCCTCGGTGACT 1648  
 DB 819 AGCAGATGTTTGGTGGATATGTTGCTGAGCTCAAGTCAATTTGCTACCTCCCGGAGACT 878  
 QY 1649 TTGCTGCAAAAGGAGAAATTTGGCAAGGAAATGTATATCATCAAGCATGGAGAAAGTCC 1708  
 DB 879 TTGTTTGAAGAAAGCGACATCGNTAGAGATGTACGTGATCAAAAGTGGTGGCGTGC 938  
 QY 1709 AAGTTCTTGGAGCCCTGTGATGTTACTTAAAGTTCTTGGTTACTTGAAGTGGGTGCGTGT 1768  
 DB 939 ANGTGTTGGAGGACCTGACACAGCATCTGTTCTGACGCTGAAGGCCGCTGCGGTNG 998  
 QY 1769 TTGGAGAAATCAGCTTCTTA-----GCAGCAGGAGGAGAAACCTCGAACTGCCAATG 1822  
 DB 999 TCGGAGAAATCAGTTTGTCTACAGTCTCTGTAANGATGANGAAGAAACAGGCGCACAGCTAATG 1058  
 QY 1823 TGTGTCGCCACCGGTT 1838  
 DB 1059 TGAAGCTCAGCGCTT 1074  
 RESULT 12  
 DN679703  
 LOCUS  
 DEFINITION  
 CGX07-F03.yld-s SHGC-CGX *Gasterosteus aculeatus* cDNA clone  
 CGX07-F03 5', mRNA sequence.  
 DN679703  
 VERSION  
 DN679703.1 GI:62021754  
 EST.  
 SOURCE  
*Gasterosteus aculeatus* (three spined stickleback)  
 ORGANISM  
*Gasterosteus aculeatus*  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; *Gasterosteus*;  
*Gasterosteus*.  
 1 (bases 1 to 1377)  
 Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,  
 Schmutz,J. and Myers,R.M.  
 TITLE  
 Expressed sequence tags from *Gasterosteus aculeatus*  
 JOURNAL  
 Unpublished (2003)  
 COMMENT  
 Contact: Grimwood, Jane  
 Stanford Human Genome Center

Stanford University School of Medicine  
975 S California Ave, Palo Alto, CA 94304, USA  
Tel: 650 320 5917  
Fax: 650 320 5801

High quality sequence stop: 816.

## FEATURES

1. 1377  
/location/qualifiers  
/organism="Gasterosteus aculeatus"  
/mol\_type="mRNA"  
/strain="Conner Creek sticklebacks, WA USA"  
/db\_xref="taxon:69223"  
/clone="CGX07-F03"  
/sex="mixed male and female"  
/tissue\_type="eyes"  
/dev\_stage="adult"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="SHGC-CGX"  
/note="Vector: Express 1; Total and poly A+ isolated from the indicated stickleback tissue library was constructed in the Express 1 plasmid. Open Biosystems. First strand cDNA synthesis with an 54 bp linker primer containing an overlap preceded by a synthetic NotI site (first 5'-GACGAGTCTAGATCGGACGCGCGCC(T)25-3'). Second strand synthesis, cDNAs were made bl corresponding to the original 5 prime end and cloned directionally into the NotI and EcoRV Express 1. Note that the EcoRV site is typically in the blunt end cloning, leaving a junction of 'xxxAATC' (where is ATC is the second half of the site, and xxx is derived from the cDNA sequence). The Express 1 vector is available at: [http://www.openbiosystems.com/cdna\\_library.q.php#8](http://www.openbiosystems.com/cdna_library.q.php#8) The primary library was transformed in DH10B (T1 phage resistant) bacteria. Clones from Open Biosystems: <http://www.openbiosystems.com/stickleback>"

## ORIGIN

Query Match	16.0%;	Score 388.8;	DB 8;	Length 1377;
Best Local Similarity	64.5%;	Pred. No. 2.4e-79;		
Matches 595; Conservative	0;	Mismatches 324;	Indels 3;	Gaps 1;
QY	1128	CTATGAAGGAATTGGGCAC	TACTAGATGGGTGTATGATGGGGAAGGAAC	CAGATATCTGAG 1187
Db	1	CTACGAAGGACTCGGATCC	ACCAAGTGGGTCTACGACGGGAAGGAAC	GCTTATATCCG 60
QY	1188	ATGTTATATTGGGCAGTT	CGAACTTTAAATTACCATTTGGTGGCTTCC	GAAACCAACAAC 1247
Db	61	CTGTTACTACTTTTGGCG	TGAAGACATTTGATCACCATCGGAGGACTT	CTCTGACCCACCAC 120
QY	1248	TTTATTTGAAATGTTTTT	CAACTCTTTGAATTTTTTTCTGGAGTTT	TGTGTTCTCCAG 1307
Db	121	CGTCTTTGAGATCTGCT	TTCGAGCTCATCAACTACTCTGTTGGGGTCT	TTCGCTTTTCGAT 180
QY	1308	TTTAAATTGGTTCAGAT	GAGAGATGTGATTTGGAGCAGCTTACAGCCAAT	CAGAACTACTTTCCG 1367
Db	181	CATGATCGGCACATGAG	AGATGTGGTTCGGAGCTCGGACCGCCGGAGAGA	ACTACTATACCG 240
QY	1368	CGCCTGCATGGATGACAC	CAATTGCGCTTACATGAACAATTTACTCATTCT	TCTAAACTTTGGCA 1427
Db	241	AGCCTCGATGGACAGCAC	CAACATCCAGTACATGAATCTTTACCAATCC	CGCGACAGGTCCA 300
QY	1428	AAAGCGAGTTTCGGACT	TGGTATGAAATATACATCGGGACTCTCAAAGAA	TGCTAGATGATC 1487
Db	301	GAACCGCATCAAGACCT	TGGTACGACTACACTTGGAGATCCAGGGCATG	CTGATGNACA 360
QY	1488	TGATTTGCTTTAGACCC	TACCACTACGGTTCAGTTAGCCCTCCGCAT	TGATGTGAACATT 1547
Db	361	GGAGCTTCTAGTGCAGCT	TTCGCGTAAAGATGAGGCTGGACATGGCTGT	AGACGTCAACTA 420

```
/tissue_type="eyes"
/dev_stage="adult"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="SHGC-CGX"
/notes="Vector: Express 1; Total and poly A+ RNA was
isolated from the indicated stickleback tissue, and a cDNA
library was constructed in the Express 1 plasmid vector by
Open Biosystems. First strand cDNA synthesis was primed
with an 54 bp linker primer containing an oligodT sequence
preceded by a synthetic NotI site (first strand primer:
5'-GACTAGTCTAGATCGGAGCGGCCCTT25-3'). Following
second strand synthesis, cDNAs were made blunt at the end
corresponding to the original 5' prime end of mRNA, and
cloned directionally into the NotI and EcoRV sites of
Express 1. Note that the EcoRV site is typically destroyed
in the blunt end cloning, leaving a junction of the form
'xxxATC' (where ATC is the second half of the EcoRV
site, and xxx is derived from the cDNA sequence). A map of
the Express 1 vector is available at:
http://www.openbiosystems.com/cdna_library_construction_fa
q.php#8 The primary library was transformed and amplified
in DH10B (T1 phage resistant) bacteria. Clones available
from Open Biosystems:
http://www.openbiosystems.com/stickleback"
```

## ORIGIN

```
Query Match 15.4%; Score 373.4; DB 8; Length 1140;
Best Local Similarity 61.5%; Pred. No. 8.9e-76;
Matches 649; Conservative 0; Mismatches 402; Indels 5; Gaps 3;
```

```
QY 617 TTCCAAACAGCATAGATTATACACAGATCGACTCTATCTCTGTGGCTCTTGTGTCA 676
DB 83 TCCCTTCCAGCATGACCGCTTCACTGATGATGATGATGATGATGATGATGATG 142
QY 677 CTCTTGCTTAATCTGAACCTGCTGTTTATACCACTGCGCTCTCTCCCATATCAA 736
DB 143 TGGGGCTGGAAGTGAAGTGGTTGATTCCTGGGCGGGGGGGCTTCCATACCAA 202
QY 737 CCGCAGACAACATCACTACTGCTGTTATTCGGACATCATATGTGATCATCTACCTTT 796
DB 203 CCCCAGAGACCTTCACTGTGGCTCTGTTGGACTACACTGTGACCTGATCTACATCA 262
QY 797 ATGATATGCTATTATCCAGCCAGACTCCAGTTTGAAGAGGAGGAGACATATAGTGG 856
DB 263 CTGACATCTGCTGATGACGCGCGCTTCACTGTTGTTGAGGAGGAGACATTTGAGCG 322
QY 857 ATTCAATGAGCTAAGGAAACACTACAGGACTTCTACAAATTTCACTGATGTCGAT 916
DB 323 ATAGAAGGACATGAGGAGACATCATATGACCACTGAGAGATTTAAGATGACGTATCA 382
QY 917 CAATAATACCAATTTGATATTGCTACCTCTTCTTTGGGTTTAAATCCAAATGTTAGACAA 976
DB 383 GCTTGTTCCTGGTGGAGATCTTTATGTTTTCCTGAGTCAACTCTCTGTTGAGGTTCC 442
QY 977 ATAGATGTTAAAGTACACTTCTTTTGAATTTATCATCATCAGTACGATCTATATGG 1036
DB 443 CTCGGCTGCTGAAGTACATGCGCTTCTTTGAGTTCAATGACCGGATGGGCTGTGATGA 502
QY 1037 ACAAGCATATATCTACAGAGTTATTCGAACTGATCTCTGTTTATTTCTGACACA 1096
DB 503 AGAAGCATATCTACAGGTTGATCCGAGCTCCAGTACCTCTCTCTCTCTGACACA 562
QY 1097 TTAATGCTGTGTTTATTTACTGGGCTTCAAACTATGAAGGAATTTGGCACTACTAGATGG 1156
DB 563 TCAAGCTGTGCTCTTTTATTTGGGATCAGACTACGAGGACTCGGATCCCAAGTGG 622
QY 1157 TGTATATGGGAGAGAAACAGATATCTGAGATGTTATTTATGGGAGTTCGAATTTAA 1216
DB 623 TCTACGAGGAGAAAGAAACGCTTATATCCGCTGTACTACTTTTGGGTGAAGACATTTGA 682
QY 1217 TTACCATTTGGTGGCTTCCAGAACACAACTTTTATTTGAATTTGTTTTCACCTTTGA 1276
DB 683 TCACCATCGGAGGACTTCTCTGACCCCAACCCGCTTTTGATGATCTGCTTCAGCTCATCA 742
```

```
QY 1277 ATTTTCTTCTGAGTCTTTTGTCTTCTCCAGTTTAAATTTGGTCAGATGAGATGTCATTC 1336
DB 743 ACTACTTCTGTTGGGCTTTTGGCTTTTTCGATCATGATCGGCGAGATGAGAGATGTGTCG 802
QY 1337 GAGCAGCTACAGCAATCAAGAACTACTTCCGGCGCTTCGATGATGATGACACCATTTGCCATCA 1396
DB 803 GAGCTGGACCGCGGAGAGAACTACTACCGAGCTGCTGATGAGACAGCACCACCTACATACA 862
QY 1397 TGAACAATTTACTTCCATTTCTTAACTTTGTGCAAAAGGAGTTCCGACTTGGTATGAATATA 1456
DB 863 TGAACCTCTTTTACACATCTCCCGGACAGGTCAGAACCGCATCAAGAACTGGTACGACTCCC 922
QY 1457 CATGGAGCTCTCAAGAAATGCTAGATGACTCTGATTTGCTTAAAGACCTTACCAACTACGG 1516
DB 923 CTTGGAAGATACAGGCGCATGCTGGATGA-ACAGAGCTTCTTAGTGCAGCTCCCGGCTAANA 981
QY 1517 TCCAGTTAGCTTCCGCTTGCATTTGATGTGAACCTTCAGCATCATCATCAAGCAAGTCTGCTT- 1575
DB 982 TGAGGCTGACCTTGGCTGGAAAACCTACACTAATCCCTTCTCAGCAAGTGGCCCTGTTTC 1041
QY 1576 AAGGCTTGTGATACACAGA---TGATTATGACATGTTGCTAAGATTGAATTCGTTCTC 1632
DB 1042 CAGGGCTGTGAAAAAATAATGGGTTTCCACCTGTTAAACCAAGTTCCAGTCAGGCTCT 1101
QY 1633 TATTTCCCTGGTGCATTTTGTCTGCAAAAAGGGAGAA 1668
DB 1102 ACCTGCCCGGAAACTTTGTCTGTAAAAAGGCGAAA 1137
```

RESULT 14  
BX401127  
LOCUS  
DEFINITION  
BX401127 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DK009YE13 5-PRIME, mRNA sequence.  
ACCESSION  
BX401127  
VERSION  
BX401127.2 GI:46846684  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1054)  
Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 13, 2003 this sequence version replaced gi:30614460.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
7817.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?se=CS0DK009AC070P1&c=7817.r.

FEATURES  
source  
1. 1054  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK009YE13"  
/call\_type="HELA CELLS COT 25-NORMALIZED"  
/cell\_line="HELA"  
/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

```
Query Match      15.0%; Score 363.8; DB 5; Length 1054;
Best Local Similarity 60.8%; Pred. No. 1.5e-73;
Matches 611; Conservative 35; Mismatches 353; Indels 6; Gaps 5;

QY 861 AAATGAGCTAAGAAACACTACAGGACTCTTACAAAATTTTCTAGTTGGATGTCGATCAAT 920
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 AAAGGATATCGGAAATAACTACGTAAGTCTCGCCGCTTCAAGATGACACTGCTCAGCCT 63

QY 921 AATACATTGATATTTGCTACTCTCTCTTTGGGTTTAAATCCAAATGTTTAGAGCAATAG 980
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 64 CTGCGCTTGGATTTTCTCTATTGAAAGTCGGTGTGAACCCCTCTCTCCGCVTGMCGG 123

QY 981 GATGTTAA--GTACACTTTCATTTTTTGAATTTAATCATCAGCTAGAGTCTATAATGGAC 1038
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 CTGTTTAAATGATCATGSCCTTCTTCGAGTTTACAGCGCGCTGGAATCCATCCTCAGC 183

QY 1039 AAAGCATATCTACAGATTTATTCGAAACAACTGGATATCTGCTGTTTATCTGCACAT 1098
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 AAAGCCTACGTGTACAGGGTCTACAGAYCACAGYTTACCTTCTTTACAGCCTGCATTG 243

QY 1099 AATGCTGTGTTTATTACTGGGCTTCACAACTATGAGGAATGGCACTACTAGATGGTG 1158
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 244 AATTCGTGCTTTTATTACTGGGCATCGGCTATCAGGGCTCGGCTCCACTCACTGGGTT 303

QY 1159 TATGATGGGGAAGAAACGAGTATCTGAGATGTTATTATTGGGCAGTTCGAACCTTTAAT 1218
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 304 TAGCATGGCGTGGGAAACAGTTATATTCGCTGTACTCTTGTGTGAACMCTCATC 363

QY 1219 ACCATTGTGGGCTTCCAGAACACAACTTTATTGAAATTTGTTTTCACACTTGAAT 1278
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 364 AYCATCGGGGGCTGCTCAGCCCAAGAYACTCTTTGAAATTTGCTTCCAGCTGVTGAAT 423

QY 1279 TTTTTTCTCGAGTTTCTGTTCTCCAGTTTAAATTTGTCAGATGAGATGTTGATGGA 1338
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 424 TATTTACGGGGCTCTTCTGTTCTGTGATGATTGGACAGATGAKAGATGTGTAGGG 483

QY 1339 GCAGCTACAGCCCAATCAGAACTACTTCCCGGCTGCTAGGATGACACCAATTCCTACATG 1398
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 484 GCGGCCACCG-CAGACAGACCTACTACCAGCTGCATGACAGCAGCAGCGTGAATGATG 542

QY 1399 ACAATTAATCTCAATCTCTTAACTTTGTGAAAAGCGAGTTTCGACTTTGGTATGAATACA 1458
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 543 AATTTTATCAAGATCCCCAAGTCCGTGCAGAAACGAGTCAAGACCTGTGTAAGATACACC 602

QY 1459 TGGGACTCTCAAGAAATGCTAGATGAGTCTGATTTGCTTTAAGACCTCAACCACTACGGTC 1518
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 603 TGGCACTGCAAGGCATGCTGGATGAGTCAGAGCTGATGGTGCAGCTTCCAGCAAGATG 662

QY 1519 CAGTTAGCCCTCGCCATTTGATGTGAATTTTCAGCATCATCAGCAAAAGTCGACTTTGTTCAAG 1578
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 663 CGGCTGGACCTCGCCATCGACATGAATACAACTCGTTAGCAAGTCGCACTCTTTGAG 722

QY 1579 GGTGTGTATACAGATGATTTATGACATGTTGCTTAAGATTTGAAATTCGGTTCTCTATTG 1638
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 723 GCGTGTGACCGGCAGATGATCTTTGACATGCTGAAGAGGCTTCGCTCTGTTGCTACCTG 782

QY 1639 CCGTGTGATTTTGTCTGCAA-AAAGGAGAAATTCGCAAGGAATGCTATATCATCAAGCA 1697
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 783 CCCAACGATATGTGTGCAGAAAGVGGGSAGATCGGCCGTGAGATGTATCATCAAGC 842

QY 1698 TGGAGAAGTCCAAGTTCTTGGAGGCCCTGATGGTACTAAAGTTCTGGTTACTCTGAAAGC 1757
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 843 AGGCAAGTGCAGGCTTGGGCGSCCTGATGGGAAATCTGTCGTGA-CCTKCAASY 901

QY 1758 TGGGTGCGTGTGAGAAATCAGCCTTCTACAGCAGGAGGAGAAACCGTCGAACTGC 1817
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 902 KSSATYTGTTGTGAGMAAAWAGATGCTGCTGCTGTGGGGCGGCAA-DGGCGCAGMGC 960

QY 1818 CAATGCTGTGGCCACGGTTTGGCCAACTCTTTTAACTCTAGACAA 1862
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 961 CAAAGTGTGSSSSMGGSSTTACMACTCTTTATCTGSAATAAA 1005
```

```
RESULT 15
DQ032774
LOCUS      3444 bp      DNA      linear      GSS 02-JUN-2005
DEFINITION Pan troglodytes CNGB1 gene, VIRTUAL TRANSCRIPT, partial sequence,
             genomic survey sequence.
ACCESSION  DQ032774
VERSION    DQ032774.1
KEYWORDS   GSS
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominoidea; Pan.
REFERENCE  1 (bases 1 to 3444)
            Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.H.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
            A Scan for Positively Selected Genes in the Genomes of Humans and
            Chimpanzees
            (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED    15869325
AUTHORS   2 (bases 1 to 3444)
            Nielsen,R., Bustamante,C., Clark,A.G., Glanowski,S., Sackton,T.H.,
            Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
            White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
            Direct Submission
            Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering
            them based on alignment. Translation starts at the beginning of
            alignment.
FEATURES   Location/Qualifiers
            source          1..3444
                        /organism="Pan troglodytes"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9598"
                        <1..>3444
                        /gene="CNGB1"
                        /locus_tag="HC2747"
ORIGIN
Query Match      13.7%; Score 333.8; DB 11; Length 3444;
Best Local Similarity 35.8%; Pred. No. 1.9e-66;
Matches 482; Conservative 0; Mismatches 865; Indels 0; Gaps 0;

QY 617 TTCAAAACAGCATAGATTACACAGATCGACTCTATCTCTCTGTGGCTCTTTGTTGCA 676
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1913 TTCCCAGAGCATTTGACCCGCTGACCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1972

QY 677 CTCCTTGCCTATAACTGGAACTGCTGGTTTATACCACCTCGGCTCGCTTCCCATATCAA 736
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1973 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2032

QY 737 CGCAGACAACATACACTACTCGGCTTATTGCGGACATCATATGATATCATCTACTCTT 796
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2033 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2092

QY 797 ATGATATGCTATTTATCCAGCCAGACTCCAGCTTTGTAAGAGGAGGAGACATAATAGTGG 856
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2093 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2152

QY 857 ATTCAAATGACTAAGGAAACACATACAGGACTTCTACAAAATTTTTCAGTTGGATGTCGCA 916
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2153 ACAAAGGACATCGGAAATAACTACTGTAANTCTCGCCGCTTCAAGNNNNNNNNNNNN 2212

QY 917 CAATAATACCAATTTGATATTTTGCTACCTCTCTCTTTGGGTTTAAATCAATGTTTAGACAA 976
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2213 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2272

QY 977 ATAGATGTTAAAGTACACTTCACTTTTGTGAATTTAATCATCATCAGCTAGAGTCTATAATGG 1036
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2005, 17:07:01 ; Search time 446 Seconds  
(without alignments)  
9684.921 Million cell updates/sec

Title: US-09-855-828-3  
Perfect score: 2430  
Sequence: 1 atgtttaaatcgtgacaaa.....aagaaaggtaagcaataa 2430

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/1/ina/1\_COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5\_COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/H\_COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/PTUS\_COMB.seq.\*  
7: /cgn2\_6/prodata/1/ina/PP\_COMB.seq.\*  
8: /cgn2\_6/prodata/1/ina/RE\_COMB.seq.\*  
9: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2049.4	84.3	2135	3	US-09-949-016-1074
2	2035.4	83.8	2139	3	US-09-949-016-5798
3	330	13.6	99370	3	US-09-949-016-12816
4	330	13.6	99370	3	US-09-949-016-17540
5	329.6	13.6	601	3	US-09-949-016-38930
6	329.6	13.6	601	3	US-09-949-016-38931
7	329.6	13.6	601	3	US-09-949-016-205286
8	329.6	13.6	601	3	US-09-949-016-205287
9	202	8.3	601	3	US-09-949-016-38898
10	202	8.3	601	3	US-09-949-016-205176
11	181.2	7.5	2796	3	US-09-949-016-5678
12	169.8	7.0	84571	3	US-09-949-016-17420
13	151.4	6.2	601	3	US-09-949-016-38897
14	151.4	6.2	601	3	US-09-949-016-205175
15	147.6	6.1	601	3	US-09-949-016-38915
16	147.6	6.1	601	3	US-09-949-016-205232
17	144.6	6.0	601	3	US-09-949-016-205228
18	144.6	6.0	601	3	US-09-949-016-205229
19	143	5.9	601	3	US-09-949-016-38929
20	143	5.9	601	3	US-09-949-016-205285
21	132.8	5.5	3470	3	US-09-949-016-4344
22	132.8	5.5	32666	3	US-09-949-016-16086
23	102.6	4.2	601	3	US-09-949-016-38918
24	102.6	4.2	601	3	US-09-949-016-205235

C 25	93.8	3.9	7218	2	US-08-232-463-14	Sequence 14, Appl
C 26	91.4	3.8	601	3	US-09-949-016-38908	Sequence 38908, A
C 27	91.4	3.8	601	3	US-09-949-016-205201	Sequence 205201, A
C 28	87	3.6	996	3	US-09-270-767-15059	Sequence 15059, A
29	85.6	3.5	2223	2	US-08-257-073-4	Sequence 4, Appli
30	79	3.3	1728	3	US-09-927-267-3	Sequence 3, Appli
31	79	3.3	2308	3	US-09-927-267-2	Sequence 2, Appli
32	79	3.3	2366	3	US-09-799-451-351	Sequence 351, App
C 33	78.8	3.2	929	3	US-09-671-317-14	Sequence 14, Appl
C 34	78	3.2	64309	3	US-09-949-016-14581	Sequence 14581, A
C 35	73.6	3.0	444	3	US-09-270-767-11944	Sequence 11944, A
C 36	73.6	3.0	7044	3	US-09-949-016-14113	Sequence 14113, A
C 37	72.8	3.0	1001	3	US-09-671-317-439	Sequence 439, App
C 38	72.6	3.0	1282	3	US-10-002-344A-89	Sequence 89, Appl
C 39	71	2.9	601	3	US-09-949-016-38919	Sequence 38919, A
C 40	71	2.9	601	3	US-09-949-016-205248	Sequence 205248, A
C 41	70	2.9	88002	3	US-09-949-002-639	Sequence 639, App
C 42	70	2.9	88002	3	US-09-949-002-717	Sequence 717, App
C 43	68.2	2.8	3117	3	US-09-614-221A-275	Sequence 275, App
C 44	67.8	2.8	4167	3	US-09-973-278-700	Sequence 700, App
C 45	66.8	2.7	612	3	US-09-902-540-1357	Sequence 1357, Ap

ALIGNMENTS

RESULT 1  
US-09-949-016-1074  
; Sequence 1074, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1074  
; LENGTH: 2135  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-949-016-1074

Query Match	84.3%	Score	2049.4	DB 3	Length	2135			
Best Local Similarity	99.2%	Pred. No. 0							
Matches	2075	Conservative	0	Mismatches	1	Indels	15	Gaps	1
Qy	340	CCACAAACAAACCGCTGCGTCTCTGTTATAATAGTATGCCGATGCCAGCTACAC	399						
Db	1	CCACAAACAAACCGCTGCGTCTCTGTTATAATAGTATGCCGATGCCAGCTACAC	60						
Qy	400	AACCTGGTGAAGAAGTGGTCAAAAGAACAGCCCTCTACAAGAAAAAGTTGGTAGAGGGA	459						
Db	61	AACCTGGTGAAGAAGTGGTCAAAAGAACAGCCCTCTACAAGAAAAAGTTGGTAGAGGGA	120						
Qy	460	GATCTCTCTCACCGAAGCCAGGCCACAACTGCAAAAGCCACGGCTGTACCACCAAGTA	519						
Db	121	GATCTCTCTCACCGAAGCCAGGCCACAACTGCAAAAGCCACGGCTGTACCACCAAGTA	180						
Qy	520	AAAGAAGCGATGATAGCCAAACAGCACTTACTACAGGCTGTTGTGGTTCAAAGCTCAAA	579						
Db	181	AAAGAAGCGATGATAGCCAAACAGCACTTACTACAGGCTGTTGTGGTTCAAAGCTCAAA	240						
Qy	580	AAGATGCCCTTTAACAGAGTACTTTAAAGCGAAATTAACCTTCCAAACAGCATAGATTTCATAC	639						



US-09-949-016-5798

Query Match		83.8%;	Score 2035.4;	DB 3;	Length 2139;	
Best Local Similarity		99.0%;	Pred. No. 0;			
Matches 2075;		Conservative	0;	Mismatches	1;	Indels 19; Gaps 2;
Qy	340	CCACAAAAACACCGCTGAGCTCTGTTATTAATAGATAGTATGCCGATGCCAGCTACAC	399			
Db	1	CCACAAAAACACCGCTGAGCTCTGTTATTAATAGATAGTATGCCGATGCCAGCTACAC	60			
Qy	400	AACCTGCTGAAGAAGATGCTCAAGAACAGCCCTCTACAGAAAAAGTTGGTAGAGGA	459			
Db	61	AACCTGCTGAAGAAGATGCTCAAGAACAGCCCTCTACAGAAAAAGTTGGTAGAGGA	120			
Qy	460	GATCTCTCTCACCCGAAGCCAGCCCAACAACTGCAAGCCCAAGCCAGCTGTACCAACAGTA	519			
Db	121	GATCTCTCTCACCCGAAGCCAGCCCAACAACTGCAAGCCCAAGCCAGCTGTACCAACAGTA	180			
Qy	520	AAAGAAAGGATGATTAAGCCAAAGAACATTAATCTACAGGCTGTGGTTCAAAAGTCAAA	579			
Db	181	AAAGAAAGGATGATTAAGCCAAAGAACATTAATCTACAGGCTGTGGTTCAAAAGTCAAA	240			
Qy	580	AGATGCTTTTAAACAGAGTACTTAAAGCGAATTAATCTTCCAAACAGCATAGATTCATAC	639			
Db	241	AGATGCTTTTAAACAGAGTACTTAAAGCGAATTAATCTTCCAAACAGCATAGATTCATAC	300			
Qy	640	ACAGATCGACTCTATCTCTGTGGCTTGTCTTGTCACTCTTGCCTTATACTGGAATGTC	699			
Db	301	ACAGATCGACTCTATCTCTGTGGCTTGTCTTGTCACTCTTGCCTTATACTGGAATGTC	360			
Qy	700	TGTTTATACCACTGGCCCTGCTTCCCATATCAAAACCGCAGACATACACTACTGG	759			
Db	361	TGTTTATACCACTGGCCCTGCTTCCCATATCAAAACCGCAGACATACACTACTGG	420			
Qy	760	CTTATTGCGACATCATATGATATCATCTACTCTTATGATATGCTATTTATCCAGCCC	819			
Db	421	CTTATTGCGACATCATATGATATCATCTACTCTTATGATATGCTATTTATCCAGCCC	480			
Qy	820	AGACTCCAGTTTGAAGAGGAGAGACATTAATAGTGGATTCAATAGAGCTTAAGGAAACAC	879			
Db	481	AGACTCCAGTTTGAAGAGGAGAGACATTAATAGTGGATTCAATAGAGCTTAAGGAAACAC	540			
Qy	880	TACAGGACTTCTACAAAATTCAGTTGGATGTGCGATCAATATACCATTTGATTTGTC	939			
Db	541	TACAGGACTTCTACAAAATTCAGTTGGATGTGCGATCAATATACCATTTGATTTGTC	600			
Qy	940	TACCTCTCTTGGTTTAACTCCAACTTTTAGACCAATAGGATGTTAAAGTACACTTCA	999			
Db	601	TACCTCTCTTGGTTTAACTCCAACTTTTAGACCAATAGGATGTTAAAGTACACTTCA	660			
Qy	1000	TTTTTTGAAATTAATCATCACTAGAGTCTATTAATGACAAAGCATATATCTACAGAGTT	1059			
Db	661	TTTTTTGAAATTAATCATCACTAGAGTCTATTAATGACAAAGCATATATCTACAGAGTT	720			
Qy	1060	ATTGGAACAACTGATATCTGCTTTTATCTGCATATTAATGCTGTGTTTATCTG	1119			
Db	721	ATTGGAACAACTGATATCTGCTTTTATCTGCATATTAATGCTGTGTTTATCTG	780			
Qy	1120	GCTTCAAACTATGAAGAAATGGCACTACTAGATGGGTGATGATGGGAAGGAAACGAG	1179			
Db	781	GCTTCAAACTATGAAGAAATGGCACTACTAGATGGGTGATGATGGGAAGGAAACGAG	840			
Qy	1180	TATCTGAGATGTTATTTATGGGAGTTTCAACTTTAATTAATGCTGTGGCTTCCAGAA	1239			
Db	841	TATCTGAGATGTTATTTATGGGAGTTTCAACTTTAATTAATGCTGTGGCTTCCAGAA	900			
Qy	1240	CCACAACTTTATTTGAAATTTTTCCTGAACTTTGAACTTTTTCCTGAGTTTGTG	1299			
Db	901	CCACAACTTTATTTGAAATTTTTCCTGAACTTTTTCCTGAGTTTGTG	960			
Qy	1300	TTCTCCAGTTTAAATTTGCTCAGATGAGATGATTTGAGCAGCTACAGCCCAATCAGAAC	1359			
Db	961	TTCTCCAGTTTAAATTTGCTCAGATGAGATGATTTGAGCAGCTACAGCCCAATCAGAAC	1020			

Qy	1360	TACTTCCGGCTGATGATGACACCATTTGCTACATGAACAAATTAATCTCCATTTCTTAA	1419
Db	1021	TACTTCCGGCTGATGATGACACCATTTGCTACATGAACAAATTAATCTCCATTTCTTAA	1080
Qy	1420	CTTGTGCAAAACGAGTTTCGGACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTA	1479
Db	1081	CTTGTGCAAAACGAGTTTCGGACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTA	1140
Qy	1480	GATGAGTCTGATTTCTTTAAGACCCCTACCACTACGGTCCAGTTAGCCCTCCCAATTTAT	1539
Db	1141	GATGAGTCTGATTTCTTTAAGACCCCTACCACTACGGTCCAGTTAGCCCTCCCAATTTAT	1200
Qy	1540	GTGAACCTTCAGCATCATCAGCAAAAGTCGACTTGTTCAG-----GCTTGTGATACACAGAT	1595
Db	1201	GTGAACCTTCAGCATCATCAGCAAAAGTCGACTTGTTCAG-----GCTTGTGATACACAGAT	1260
Qy	1596	GATTTATGACATGTTGCTAAGATTGAAATCCGTTCTCTATTTGCTGCTGGTACTTTGCTG	1655
Db	1261	GATTTATGACATGTTGCTAAGATTGAAATCCGTTCTCTATTTGCTGCTGGTACTTTGCTG	1320
Qy	1656	CAAAAAGGAGAAATTTGCAAGAAATGTATATCATCAAGCATGGAAGTCCAAGTTCT	1715
Db	1321	CAAAAAGGAGAAATTTGCAAGAAATGTATATCATCAAGCATGGAAGTCCAAGTTCT	1380
Qy	1716	TGAGGCCCTGATGTTACTTAAAGTTCTGGTTACTCTGAAAGCTGGGTGGTGTGGAGA	1775
Db	1381	TGAGGCCCTGATGTTACTTAAAGTTCTGGTTACTCTGAAAGCTGGGTGGTGTGGAGA	1431
Qy	1776	AATCAGCTTCTAGCAGCAGGAGGAGAAACCGTCGAATGCGAATGTCGCTGGTGGCCACCG	1835
Db	1432	-----CCTTTAGCAGCAGGAGGAGAAACCGTCGAATGCGAATGTCGCTGGTGGCCACCG	1485
Qy	1836	GTTTGGCAATCTTTTAACTCTAGACAAAAAGACCCCTCCAAGAAATTTCTAGTGCATTTATCC	1895
Db	1486	GTTTGGCAATCTTTTAACTCTAGACAAAAAGACCCCTCCAAGAAATTTCTAGTGCATTTATCC	1545
Qy	1896	AGATTCGAAAGATCCTCATGAAGAAAGCCAGAGTGTCTTTTAAAGCAGAGGCTAAGAC	1955
Db	1546	AGATTCGAAAGATCCTCATGAAGAAAGCCAGAGTGTCTTTTAAAGCAGAGGCTAAGAC	1605
Qy	1956	CGCAGAAACACCCCTCCAAAGAAAGATCTTGCCCTCTCTTCCACCCGAAAGAGAGAC	2015
Db	1606	CGCAGAAACACCCCTCCAAAGAAAGATCTTGCCCTCTCTTCCACCCGAAAGAGAGAC	1665
Qy	2016	ACCCAACTGTTTAAACTCTCCTAGGAGGCACAGGAAAGCAAGTCTTTCAGAGCTACT	2075
Db	1666	ACCCAACTGTTTAAACTCTCCTAGGAGGCACAGGAAAGCAAGTCTTTCAGAGCTACT	1725
Qy	2076	CAAAATTTGAAGCAGAGCAAGCTCAGAAAGAAAGAAATTTCTGAAGGAGGAGAGAGA	2135
Db	1726	CAAAATTTGAAGCAGAGCAAGCTCAGAAAGAAAGAAATTTCTGAAGGAGGAGAGAGA	1785
Qy	2136	AGAAAAGAAATGAAGATAAACAAGAAAGAAATGAAGATAAACAAGAAAGAAATGAAGA	2195
Db	1786	AGAAAAGAAATGAAGATAAACAAGAAAGAAATGAAGATAAACAAGAAAGAAATGAAGA	1845
Qy	2196	TAAAGAAAGAAATGAAGATAAAGATAAAGAAAGAGAGCCAGAAAGAGAGCCACTGGA	2255
Db	1846	TAAAGAAAGAAATGAAGATAAAGATAAAGAAAGAGAGCCAGAAAGAGAGCCACTGGA	1905
Qy	2256	CAGACCTGAAATGTACAGCAAGTCTTATGTCAGTGGAGGAGAAACCCCACTCAGTTAGAG	2315
Db	1906	CAGACCTGAAATGTACAGCAAGTCTTATGTCAGTGGAGGAGAAACCCCACTCAGTTAGAG	1965
Qy	2316	GACAGTTTACCCAGAGGAGCTTCTGTCATCACTCATTTATCAGCATGGCTCTCTCTGC	2375
Db	1966	GACAGTTTACCCAGAGGAGCTTCTGTCATCACTCATTTATCAGCATGGCTCTCTCTGC	2025
Qy	2376	TGAGGGCGGAGAGAGGTTCTTACTATTGAAGTCAAAAGAAAGGCTTAAGCAATAA	2430
Db	2026	TGAGGGCGGAGAGAGGTTCTTACTATTGAAGTCAAAAGAAAGGCTTAAGCAATAA	2080







Qy	833	TAAGAGGAGGAGACATATAGTGGATTCAATGACCTAAGGNAACACTACAGGACTTCTA	892
Db	80762	TAGAACCAAGGACTGCTGGTAAAGSAGAACTTTAAACTCATAAATAAATCAAACT	80821
Qy	893	CAAAATTTTCAGTTGGATGCGCATCAATAATACCATTTGCATATTGCTACCTCTCTTTG	952
Db	80822	TGCAAATTTAAACTTGATGTTCTGTCACATACCAACTGATTTGCTGATTTTAAAGTTAG	80881
Qy	953	GGTTTA---ATCCAATGTTTAGAGCAAAATAGAGATGTTAAAGTACACTTTCATTTTTGAAT	1009
Db	80882	GGTGGAACTATCCAGAAATTAGATTAAACAGGTGTTACGGTTCCTCGATGTTTGGAT	80941
Qy	1010	TTAATCATCACTTAGAGTCTATAATGGCAAAAGCATATATACAGAGTTATTTCGAACAA	1069
Db	80942	TC TTCAGAAACAGAAACAAGCAAACTATCCAAACATCTTCAGGATTTTCCAACCTTG	81001
Qy	1070	CTGGATCTTGCTGTTTTATTCTGCACATTTAAATGCCTGTGTTTTATTACTGGGCTTCAAACT	1129
Db	81002	TTATGTATATCGTCATCATTTATCCACTGGAATGCATGTGTGTTCTACTCTATTTCTTAAG	81061

```
QY 1130 ATGAAGGATTTGGCACTACTAGATGGGTGTA-----TGATG 1165
Db 81062 CTATTGGATTGGAAATGATACATAGGCTCTACCCCTGATATTAATGATCCTGAATTTGGCC 81121
QY 1166 GGGAGGAAACAGATATCTGAGATGTTATTATTGGGCAGTTCGAACCTTTAAATTAACCATTTG 1225
Db 81122 GTTTGGCTAGAAATACGTATACAGCCTTTACTGGTCTACATGACTTTTGACTACCATTTG 81181
QY 1226 GTGGCCTTCCAGAACCAACAACTTTATTGAAATTTGTTTTTCAACTCTTTGAAATTTTTTTT 1285
Db 81182 GTGAACACCCCTCCCGTGAGGATTCGTGAGTATGCTTTGTGGTGGTTGATTTCTCTAA 81241
QY 1286 CTGGAGTTTGTGTTCTCCAGTTTAATTTGGTCAGATGAGAGATGATGATTTGAGCAGCTA 1345
Db 81242 TTGGAGTGTAAATTTTGTACCATCGTTGGTAACATAGTGTCTATGATTTTCCAAACATCA 81301
QY 1346 CAGCCAAATCAGAACTACTTTCGGCGCTGCATGATGATGACACCACTTGCTCATACACAAAT 1405
Db 81302 ATGCAGCCAGAGCAGAAATTTCAAGCAAGAAATGATGCTATCAAGCAATATATGCAATTTT 81361
QY 1406 ACTCCATTCTAAACTTTGTGCAAAAGCGAGTTGGGACTTTGGTATGAAATATACATGGGACT 1465
Db 81362 GAAATGTAAGCAAGATATATGGAAGAGAGGTTATTAAATGGTTTGACTACCTGTGGACCA 81421
QY 1466 CTCGAAGATGCTAGATGATCTGATTTGCTTTAAGACCCCTACCAACTACGGTCCAGTTAG 1525
Db 81422 ACAAATAAAACAGTTGATGAGAAAGAGTCTTAAAGTATCTACCTGATATAAATAAGAGCAG 81481
QY 1526 CCCTGCCATGATGTGAATCTTCAGCATCATCAGCAAAAGTCGACTTTGTTCAAGGGTTGG 1585
Db 81482 AAATTGCCATCAACGTTTCACTTAGACACATTTAAAGAGTACGCAATTTTGTGCTGATTTG 81541
QY 1586 ATACACAGATGATTTATGACATGTTGCTAAGATTTGAAATCCGTTCTCTATTTGCTGTTG 1645
Db 81542 AAGCTGCTGTTTGGTGGAGTTGGTCTTGAATTTGCAACCCCAAGTCTACAGTCTCTGGAG 81601
QY 1646 ACTTTGCTGCAAAAGGAGAGAAATTTGGCAAGAAATGATATCATCAAGCATGAGGAAG 1705
Db 81602 ATTATATTTGCAAGAAAGGGGATATCGGACGAGAGATGTACATTTATCAAGGAAGGCAAC 81661
QY 1706 TCCAGTTCTTGAGGCCCTGATGGTACTAAAGTTCTGCTGTTTCTGCTGAAAGCTGGG 1761
Db 81662 TGCTGTGCTGCGAGATGATGAGTCACTCAGTTGTGTTATTTGAGCGATGCGCACTACT 81721
QY 1762 TCGGTGTTTGGAGAAATC--AGCCTTCTAGCAGCAGGAGGAGAAACCGTCGAACTGCCA 1819
Db 81722 TCGGTGAGATCAGCATTTCTTAACATTTAAAGGAGCAAAAGCTGCAATCGAAGAACGGCCA 81781
QY 1820 ATGTGTTGCCCCACGGTTTGGCAATCTTTAACTCTAGACAAAGAACCCCTCCAGAA 1879
Db 81782 ATATTAAAGATATTGGCTACTCAGACCTGTTCTGCTCTCTCAAAAGATGACCTCATGGAAG 81841
QY 1880 TTCTAGTCATTATCCAGATTTCTGAAGGATCCTCATGAAGAAAGCCAGAGTCTTTTAA 1939
Db 81842 CTCTAACTAGTACCCAGATGCAAAACTATGCTGGAAGAAAGGGAAGAGAGATTTTAA 81901
QY 1940 ACAGAAGGCT 1950
Db 81902 TGAAGATGGT 81912
```

```
RESULT 13
US-09-949-016-38897/c
; Sequence 38897, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38897
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-38897

Query Match 6.2%; Score 151.4; DB 3; Length 601;
Best Local Similarity 99.3%; Pred. No. 2.3e-27;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 493 GCAAGCCACGGCTGTACCAACAGTAAAGAAAGCGATGATAGCCCAACAGCAATTAC 552
Db 291 GCAAGCCACGGCTGTACCAACAGTAAAGAAAGCGATGATAGCCCAACAGCAATTAC 232
QY 553 TACAGGCTGTTGTGGTTCAAAAGTCAAAAAGATGCCCTTTAAACAGAGTACTTTAAAGCGAATT 612
Db 231 TACAGGCTGTTGTGGTTCAAAAGTCAAAAAGATGCCCTTTAAACAGAGTACTTTAAAGCGAATT 172
QY 613 AAACCTTCCAAACAGCATAGATTCATACACAGAT 645
Db 171 AAACCTTCCAAACAGCATAGATTCATACACAGAT 139
```

```
RESULT 14
US-09-949-016-205175/c
; Sequence 205175, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 205175
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-205175
```

```
Query Match 6.2%; Score 151.4; DB 3; Length 601;
Best Local Similarity 99.3%; Pred. No. 2.3e-27;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 493 GCAAGCCACGGCTGTACCAACAGTAAAGAAAGCGATGATAGCCCAACAGCAATTAC 552
Db 291 GCAAGCCACGGCTGTACCAACAGTAAAGAAAGCGATGATAGCCCAACAGCAATTAC 232
QY 553 TACAGGCTGTTGTGGTTCAAAAGTCAAAAAGATGCCCTTTAAACAGAGTACTTTAAAGCGAATT 612
Db 231 TACAGGCTGTTGTGGTTCAAAAGTCAAAAAGATGCCCTTTAAACAGAGTACTTTAAAGCGAATT 172
QY 613 AAACCTTCCAAACAGCATAGATTCATACACAGAT 645
Db 171 AAACCTTCCAAACAGCATAGATTCATACACAGAT 139

RESULT 15
```

```

US-09-949-016-38915/c
; Sequence 38915, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38915
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-38915

```

```

Query Match      6.1%; Score 147.6; DB 3; Length 601;
Best Local Similarity 97.4%; Pred. No. 2e-26;
Matches 150; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1336 GGAGCAGCTACAGCAATCAGAACTACTTCGGCGCCTGCGATGATGACACCATTTGCCTAC 1395
Db 601 GGAGCAGCTACAGCAATCAGAACTACTTCGGCGCCTGCGATGATGACACCATTTGCCTAC 542

QY 1396 ATGAACAATTACTCCATTCCTAACTTTGTGCAAAAGCGAGTTCGGACTTGGTATGAATAT 1455
Db 541 ATGAACAATTACTCCATTCCTAACTTTGTGCAAAAGCGAGTTCGGACTTGGTATGAATAT 482

QY 1456 ACATGGGACTCTCAAGAATGCTAGATGAGTCTG 1489
Db 481 ACATGGGACTCTCAAGAATGCTAGGTAAGCATG 448

```

Search completed: December 23, 2005, 23:22:17  
Job time : 451 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2005, 17:11:54 ; Search time 1732 Seconds  
(without alignments)  
11601.956 Million cell updates/sec

Title: US-09-855-828-3  
Perfect score: 2430  
Sequence: 1 agttaaatacgtgacaaa.....aagaaagcgtaagcaataa 2430

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2430	100.0	2430	3	US-09-855-828-3
2	2430	100.0	2703	3	US-09-855-828-2
3	2430	100.0	2757	3	US-09-855-828-16
4	1316	54.2	1400	10	US-11-060-756-2833
5	1316	54.2	1400	10	US-11-060-756-7105
6	654.4	26.9	2516	9	US-10-450-763-5457
7	577.4	23.8	2607	6	US-10-189-507-3
8	577.4	23.8	4382	6	US-10-159-563-147
9	573.6	23.6	3811	9	US-10-450-763-27462
10	518	21.3	2244	9	US-10-450-763-11980
11	518	21.3	2244	9	US-10-450-763-11988
12	306	12.6	680	5	US-10-027-632-204539
13	306	12.6	680	6	US-10-027-632-204539
14	223	9.2	3275	10	US-11-097-143-26426
15	176.4	7.3	2500	6	US-10-172-118-427
16	176.4	7.3	2500	7	US-10-342-887-427
17	165	6.8	581	4	US-09-925-065A-485501
18	160.6	6.6	289	9	US-10-450-763-11095
19	160	6.6	563	4	US-09-925-065A-284517
20	160	6.6	1127	4	US-09-925-065A-284516
21	136.2	5.6	2109	10	US-11-097-143-24368
22	133.2	5.5	3027	6	US-10-295-573-1
23	132.8	5.5	2085	6	US-10-345-680-27

24	132.8	5.5	3486	6	US-10-345-680-25	Sequence 25, Appli
25	131.6	5.4	1995	5	US-10-087-217-1	Sequence 1, Appli
26	131.6	5.4	1995	5	US-10-087-217-5	Sequence 5, Appli
27	131.6	5.4	1995	9	US-10-978-282-1	Sequence 1, Appli
28	131.6	5.4	1995	6	US-10-978-282-5	Sequence 2, Appli
29	131.6	5.4	3027	6	US-10-295-573-2	Sequence 2, Appli
30	131.6	5.4	3027	6	US-10-295-573-4	Sequence 4, Appli
31	129	5.3	526	3	US-09-864-761-12975	Sequence 12975, A
32	128.4	5.3	1995	5	US-10-087-217-3	Sequence 3, Appli
33	128.4	5.3	1995	5	US-10-087-217-7	Sequence 7, Appli
34	128.4	5.3	1995	9	US-10-978-282-3	Sequence 3, Appli
35	128.4	5.3	1995	9	US-10-978-282-7	Sequence 7, Appli
36	128	5.3	2877	6	US-10-295-573-3	Sequence 3, Appli
37	127	5.2	127	3	US-09-864-761-29538	Sequence 29538, A
38	124.4	5.1	5133	10	US-11-097-143-19637	Sequence 19637, A
39	120.6	5.0	5692	10	US-11-097-143-26425	Sequence 26425, A
40	111.2	4.6	1995	3	US-09-735-927-1	Sequence 1, Appli
41	111.2	4.6	1995	5	US-10-034-843-1	Sequence 1, Appli
42	111.2	4.6	1995	6	US-10-168-651-34	Sequence 34, Appli
43	111.2	4.6	1995	6	US-10-189-507-1	Sequence 1, Appli
44	111.2	4.6	2111	6	US-10-114-153-17	Sequence 17, Appli
45	111.2	4.6	2190	5	US-10-029-677-23	Sequence 23, Appli

ALIGNMENTS

RESULT 1  
US-09-855-828-3  
; Sequence 3, Application US/09855828  
; Publication No. US20040137433A1  
; GENERAL INFORMATION:  
; APPLICANT: Cretech, Christopher D.  
; APPLICANT: Jegla, Timothy J.  
; APPLICANT: ICAGEN, Inc.  
; TITLE OF INVENTION: CNG3B: A Novel Cyclic Nucleotide-Gated Cation Channel  
; FILE REFERENCE: 018512-006010US  
; CURRENT APPLICATION NUMBER: US/09/855,828  
; PRIOR FILING DATE: 2001-05-14  
; PRIOR FILING DATE: 2000-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2430  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human CNG3B coding sequence  
; NAME/KEY: CDS  
; LOCATION: (1)..(2430)  
; OTHER INFORMATION: CNG3B  
US-09-855-828-3

Query Match	100.0%;	Score	2430;	DB	3;	Length	2430;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	2430;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	ATGTTTAAATCGTGACAAAAGTCAACAAGTCAAGGCTATAGAGAGAAACAATGAGAAT	60				
Db	1	ATGTTTAAATCGTGACAAAAGTCAACAAGTCAAGGCTATAGAGAGAAACAATGAGAAT	60				
Qy	61	GAACAAAGTTCTCGTCGGAATGAAGAGGCTCTCACCAAGTAATCAAGTCTCAGCAAAACC	120				
Db	61	GAACAAAGTTCTCGTCGGAATGAAGAGGCTCTCACCAAGTAATCAAGTCTCAGCAAAACC	120				
Qy	121	ACAGCAG	180				
Db	121	ACAGCAG	180				
Qy	181	ACGTTCTGAAGGCGCACACACACATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	240				
Db	181	ACGTTCTGAAGGCGCACACACACATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	240				

Db 181 ACCTCTGAAGAGCCACACACACCAACATACAAGACAAAACTCTCCAAGAAAAAATTCCTCTGGA 240  
Qy 241 GATCTGACCAAAACCCCTGACCCCTCAAAATGCAGCAGAAACCAACTGGAAACAGTGCAGAG 300  
Db 241 GATCTGACCAAAACCCCTGACCCCTCAAAATGCAGCAGAAACCAACTGGAAACAGTGCAGAG 300  
Qy 301 CAGAGGAAATGGACCCCGGGGAAAGAGTGCACACAGCCCAAAACCAAAACCGCTGCA 360  
Db 301 CAGAGGAAATGGACCCCGGGGAAAGAGTGCACACAGCCCAAAACCAAAACCGCTGCA 360  
Qy 361 GCTCTCTGTTATAATGAGTATGCCGATGCCAGCTACACAACTGGTGAAGAGATGCGT 420  
Db 361 GCTCTCTGTTATAATGAGTATGCCGATGCCAGCTACACAACTGGTGAAGAGATGCGT 420  
Qy 421 CAAAGAACAGCCCTCTCAAGAAAAAGTTGGTAGAGGGAGATCTCTCTCACCCGGAAGCC 480  
Db 421 CAAAGAACAGCCCTCTCAAGAAAAAGTTGGTAGAGGGAGATCTCTCTCACCCGGAAGCC 480  
Qy 481 AGCCCAAACTGCAAAAGCCACGGCTGTACCAACAGTAAAGAAAGCGATGATAGCCA 540  
Db 481 AGCCCAAACTGCAAAAGCCACGGCTGTACCAACAGTAAAGAAAGCGATGATAGCCA 540  
Qy 541 ACAGAACATTACTACAGGCTGTGTGGTTCNAAGTCAAAAGATGSCCTTTAACAAGATAC 600  
Db 541 ACAGAACATTACTACAGGCTGTGTGGTTCNAAGTCAAAAGATGSCCTTTAACAAGATAC 600  
Qy 601 TTAAGCGAATTAACCTTCAAAACAGCATAGATTTCATACACAGATCGACTCTATCTCCTG 660  
Db 601 TTAAGCGAATTAACCTTCAAAACAGCATAGATTTCATACACAGATCGACTCTATCTCCTG 660  
Qy 661 TGGCTCTTGGTGTACTCTTGGCTATAACTGGAAGTGTGGTTATACCACTGGCCCTC 720  
Db 661 TGGCTCTTGGTGTACTCTTGGCTATAACTGGAAGTGTGGTTATACCACTGGCCCTC 720  
Qy 721 GTCTTCCCATACAAACCCGACACATACACTACTGCTTATTCGAGACATCATATGT 780  
Db 721 GTCTTCCCATACAAACCCGACACATACACTACTGCTTATTCGAGACATCATATGT 780  
Qy 781 GATATCATCTACTCTTATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAAAGGA 840  
Db 781 GATATCATCTACTCTTATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAAAGGA 840  
Qy 841 GGAGACATAATAGTGAATTCAAATGAGCTAAGGAAACACTACAGGACTTCTACAAAATTT 900  
Db 841 GGAGACATAATAGTGAATTCAAATGAGCTAAGGAAACACTACAGGACTTCTACAAAATTT 900  
Qy 901 CAGTTGGATGTCGCATCAATATACCATTTGATATTTGCTACTCTCTCTTTGGGTTTAT 960  
Db 901 CAGTTGGATGTCGCATCAATATACCATTTGATATTTGCTACTCTCTCTTTGGGTTTAT 960  
Qy 961 CCAATGTTTAGAGCAAAATAGGATGTTAAAGTACACTTCAATTTTGAATTTAATCATCAC 1020  
Db 961 CCAATGTTTAGAGCAAAATAGGATGTTAAAGTACACTTCAATTTTGAATTTAATCATCAC 1020  
Qy 1021 CTAGAGTCTATTAATGGACAAAGCATATATCTACAGATTAATTCGAACAACTGGATCTTG 1080  
Db 1021 CTAGAGTCTATTAATGGACAAAGCATATATCTACAGATTAATTCGAACAACTGGATCTTG 1080  
Qy 1081 CTGTTTATCTGCACATTAATGCTGTGTTTATTTACTGGGCTTCAAACTATGAAGGAAT 1140  
Db 1081 CTGTTTATCTGCACATTAATGCTGTGTTTATTTACTGGGCTTCAAACTATGAAGGAAT 1140  
Qy 1141 GGCACCTACTAGATGGGTATGATGGGGAAGGAAACGAGTATCTGAGATGTTTATTATGG 1200  
Db 1141 GGCACCTACTAGATGGGTATGATGGGGAAGGAAACGAGTATCTGAGATGTTTATTATGG 1200  
Qy 1201 GCAGTTCGAACCTTTAATTAACATTTGGTGGCTTTCAGAACCAAACTTTATTTGAAAT 1260  
Db 1201 GCAGTTCGAACCTTTAATTAACATTTGGTGGCTTTCAGAACCAAACTTTATTTGAAAT 1260  
Qy 1261 GTTTTTCAACTCTTGAATTTTTTTCTGGAGTTTTTGTCTCCAGTTTATTTGGTCAG 1320  
Db 1261 GTTTTTCAACTCTTGAATTTTTTTCTGGAGTTTTTGTCTCCAGTTTATTTGGTCAG 1320

Qy 1321 ATGAGAGATGTGATTTGGAGCAGCTACAGCCAATCAGAACTACTTCCGCGCCTGCATGGAT 1380  
Db 1321 ATGAGAGATGTGATTTGGAGCAGCTACAGCCAATCAGAACTACTTCCGCGCCTGCATGGAT 1380  
Qy 1381 GACACATTTGCTTACATGAACAAATTAATCTTCAATTTCTTAACTTTGTGCAAAAGCGAGTTGG 1440  
Db 1381 GACACATTTGCTTACATGAACAAATTAATCTTCAATTTCTTAACTTTGTGCAAAAGCGAGTTGG 1440  
Qy 1441 ACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTAGATGAGTCTGATTTGGCTTAAG 1500  
Db 1441 ACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTAGATGAGTCTGATTTGGCTTAAG 1500  
Qy 1501 ACCCTACCAACTACCGTCCAGTCCAGTCCGCTAGCTTGTCTGCAAAAGGGAGAAATTCGTAAAG 1560  
Db 1501 ACCCTACCAACTACCGTCCAGTCCAGTCCGCTAGCTTGTCTGCAAAAGGGAGAAATTCGTAAAG 1560  
Qy 1561 AAAGTCGACTTTGTTCAAGGGTTGTATACACAGATGATTTATGACATGTTGTTAAAGATTG 1620  
Db 1561 AAAGTCGACTTTGTTCAAGGGTTGTATACACAGATGATTTATGACATGTTGTTAAAGATTG 1620  
Qy 1621 AAATCCGTTCTCTATTTTCCCTGGTGACTTTGTCTGCAAAAGGGAGAAATTCGTAAAG 1680  
Db 1621 AAATCCGTTCTCTATTTTCCCTGGTGACTTTGTCTGCAAAAGGGAGAAATTCGTAAAG 1680  
Qy 1681 ATGTATATCATCAAGCATGGAGAGTCCCAAGTTCTTGAGGCCCTGTGTTACTTAAAGTT 1740  
Db 1681 ATGTATATCATCAAGCATGGAGAGTCCCAAGTTCTTGAGGCCCTGTGTTACTTAAAGTT 1740  
Qy 1741 CTGGTTACTCTGAAAGCTGGGTGGTGGAGAAATCAAGCTTCTAGCAGCAGGAGGA 1800  
Db 1741 CTGGTTACTCTGAAAGCTGGGTGGTGGAGAAATCAAGCTTCTAGCAGCAGGAGGA 1800  
Qy 1801 GGAAACCGTCCGAACTGCTCAATGCTGGTGGCCACCGGTTTGCCTTAACTCTAGAC 1860  
Db 1801 GGAAACCGTCCGAACTGCTCAATGCTGGTGGCCACCGGTTTGCCTTAACTCTAGAC 1860  
Qy 1861 AAAAGAGCCCTCCAAGAAATTTCTAGTGCATTTATCCAGATTTCTGAAAGGATCCTCATGAAG 1920  
Db 1861 AAAAGAGCCCTCCAAGAAATTTCTAGTGCATTTATCCAGATTTCTGAAAGGATCCTCATGAAG 1920  
Qy 1921 AAAGCAGAGTGTCTTTTAAAGCAGAGGCTAAGACCGCAGAGCAACCCCTCCAAGAAAA 1980  
Db 1921 AAAGCAGAGTGTCTTTTAAAGCAGAGGCTAAGACCGCAGAGCAACCCCTCCAAGAAAA 1980  
Qy 1981 GATCTTGGCCCTCCTCTTCCACCGAAAGAGACACCCAACTGTTTTAAACTCTCCTA 2040  
Db 1981 GATCTTGGCCCTCCTCTTCCACCGAAAGAGACACCCAACTGTTTTAAACTCTCCTA 2040  
Qy 2041 GGAGCACAAGGAAAGCAAGTCTTGGCAAGACTACTCAAAATTTGAAGCAGAGCAAGCAGCT 2100  
Db 2041 GGAGCACAAGGAAAGCAAGTCTTGGCAAGACTACTCAAAATTTGAAGCAGAGCAAGCAGCT 2100  
Qy 2101 CAGAAGAAAGAAAAATTTCTGAAGGAGGAGAGAGAGAAAAAATGAAGATAAACA 2160  
Db 2101 CAGAAGAAAGAAAAATTTCTGAAGGAGGAGAGAGAGAAAAAATGAAGATAAACA 2160  
Qy 2161 AAAGAAATGAAGATAAACAAGAAAAAATCAAGATAAGGAAAGAAAAATCAAGATAA 2220  
Db 2161 AAAGAAATGAAGATAAACAAGAAAAAATCAAGATAAGGAAAGAAAAATCAAGATAA 2220  
Qy 2221 GATAAGGAAGAGAGCCAGAAAGCAAGCTCTGACAGACCTGAATGTACAGCAAGTCTT 2280  
Db 2221 GATAAGGAAGAGAGCCAGAAAGCAAGCTCTGACAGACCTGAATGTACAGCAAGTCTT 2280  
Qy 2281 ATTGAGTGGAGGAGAAACCCCACTCAGTTTGAAGGACAGTTTTTACCCAGAGGAAATTTCT 2340  
Db 2281 ATTGAGTGGAGGAGAAACCCCACTCAGTTTGAAGGACAGTTTTTACCCAGAGGAAATTTCT 2340  
Qy 2341 CGTCAATCACTCATTTATCAGCATGCTCTCTCTGAGGCGGAGAGAGGTTCTTACT 2400  
Db 2341 CGTCAATCACTCATTTATCAGCATGCTCTCTCTGAGGCGGAGAGAGGTTCTTACT 2400



Db 1683 AAATCGTCTCTATTGGCTGGTGAATTTGTCTGCAAAAGGGAGANAATGGCAAGGAA 1742  
Qy 1681 ATGTATATCATCAAGCATGAGAAAGTCCAAAGTCTTGGAGGCCCTCATGTACTAAAGTT 1740  
Db 1743 ATGTATATCATCAAGCATGAGAAAGTCCAAAGTCTTGGAGGCCCTCATGTACTAAAGTT 1802  
Qy 1741 CTGGTTACTCTGAAAGCTGGGTGGTGGTTTGGAGAAATCAGCCTTCTAGCAGCAGGAGGA 1800  
Db 1803 CTGGTTACTCTGAAAGCTGGGTGGTGGTTTGGAGAAATCAGCCTTCTAGCAGCAGGAGGA 1862  
Qy 1801 GGAAACCGTCCAACTCCCAATGCTGGTGGCCACGGGTTTGGCCAACTCTTAACTCTAGAC 1860  
Db 1863 GGAAACCGTCCAACTCCCAATGCTGGTGGCCACGGGTTTGGCCAACTCTTAACTCTAGAC 1922  
Qy 1861 AAAAGACCTCCCAAGAAATCTTAGTGCATTTATCCAGATTTCTGAAAGGATCTCTATGAAG 1920  
Db 1923 AAAAGACCTCCCAAGAAATCTTAGTGCATTTATCCAGATTTCTGAAAGGATCTCTATGAAG 1982  
Qy 1921 AAAGCCAGAGTGTCTTTAAAGCAGAAAGGCTTAAGACCCGAGAAACCCCTCCAAAGAAA 1980  
Db 1983 AAAGCCAGAGTGTCTTTAAAGCAGAAAGGCTTAAGACCCGAGAAACCCCTCCAAAGAAA 2042  
Qy 1981 GATCTTGCCCTCTCTTCCACCGAAAGAGAGACACCCAACTGTTTAAACTCTCCTA 2040  
Db 2043 GATCTTGCCCTCTCTTCCACCGAAAGAGAGACACCCAACTGTTTAAACTCTCCTA 2102  
Qy 2041 GGAGGCACAGGAAAGCAAGTCTTGAAGACTACTCAAAATTTGAAGGAGCAAGCAGCT 2100  
Db 2103 GGAGGCACAGGAAAGCAAGTCTTGAAGACTACTCAAAATTTGAAGGAGCAAGCAGCT 2162  
Qy 2101 CAGAAGAAAGAAATTTCTGAAGGAGAGAGAGAGAAAGGAAAGAAATGAAGATAAA 2160  
Db 2163 CAGAAGAAAGAAATTTCTGAAGGAGAGAGAGAGAAAGGAAAGAAATGAAGATAAA 2222  
Qy 2161 AAAGAAATGAAGATAAAAGAAAGAAATGAAGATAAAAGGAAAGAAATGAAGATAAA 2220  
Db 2223 AAAGAAATGAAGATAAAAGAAAGAAATGAAGATAAAAGGAAAGAAATGAAGATAAA 2282  
Qy 2221 GATAAGGAG 2280  
Db 2283 GATAAGGAG 2342  
Qy 2281 ATTGCAGTGGAG 2340  
Db 2343 ATTGCAGTGGAG 2402  
Qy 2341 CGTCAATCACTCATTTATCATGATGCTCTCTCTGCTGAGGGCGGAGAGAGGTTCTTACT 2400  
Db 2403 CGTCAATCACTCATTTATCATGATGCTCTCTCTGCTGAGGGCGGAGAGAGGTTCTTACT 2462  
Qy 2401 ATTGAAGTCAAGAAAGGCTTAAGCAATAA 2430  
Db 2463 ATTGAAGTCAAGAAAGGCTTAAGCAATAA 2492

RESULT 3  
US-09-855-828-16  
; Sequence 16, Application US/09855828  
; Publication No. US20040137433A1  
; GENERAL INFORMATION:  
; APPLICANT: Creech, Christopher D.  
; APPLICANT: Jegia, Timothy J.  
; APPLICANT: ICAGEN, Inc.  
; TITLE OF INVENTION: CNG3B: A Novel Cyclic Nucleotide-Gated Cation Channel  
; FILE REFERENCE: 018512-006010US  
; CURRENT APPLICATION NUMBER: US/09/855,828  
; CURRENT FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: US 60/204,445  
; PRIOR FILING DATE: 2000-05-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16

; LENGTH: 2757  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: complete CNG3B sequence derived from assembly of  
; OTHER INFORMATION: PCR fragments  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (112)..(2541)  
; OTHER INFORMATION: CNG3B  
US-09-855-828-16

Query Match 100.0%; Score 2430; DB 3; Length 2757;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2430; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGTTTAAATCGCTGACAAAAAGTCAAAAGGTGAAGCCTATATAGGAGAGAAACAATCAGAA 60  
Db 112 ATGTTTAAATCGCTGACAAAAAGTCAAAAGGTGAAGCCTATATAGGAGAGAAACAATCAGAA 171  
Qy 61 GAAACAAAGTCTCGTCGGAATGAAGAAGCCTCTCACCCAAAGTAATCAGTCTCAGCAAAACC 120  
Db 172 GAAACAAAGTCTCGTCGGAATGAAGAAGCCTCTCACCCAAAGTAATCAGTCTCAGCAAAACC 231  
Qy 121 ACAGCACAGGAGAGAAAAAAGGTTGAAGAGAAATCTCTCAAAACCAAGTCAATCTCAATTC 180  
Db 232 ACAGCACAGGAGAGAAAAAAGGTTGAAGAGAAATCTCTCAAAACCAAGTCAATCTCAATTC 291  
Qy 181 AGCTCTGAAGAGCCACACCAACATACAAAGCAAACTCTCCNAGAAAAATTTCTCTCGA 240  
Db 292 AGCTCTGAAGAGCCACACCAACATACAAAGCAAACTCTCCNAGAAAAATTTCTCTCGA 351  
Qy 241 GATCTGACACAAACCCCTGACCTCAAAATGTCAGCAGAACCAACTGGAAACAGTGGCCAGAG 300  
Db 352 GATCTGACACAAACCCCTGACCTCAAAATGTCAGCAGAACCAACTGGAAACAGTGGCCAGAG 411  
Qy 301 CAGAAAGGAAATGGACCCCGGGAAGAGAGGTCAAAACAGCCCAACAAACAAACCCGCTGCA 360  
Db 412 CAGAAAGGAAATGGACCCCGGGAAGAGAGGTCAAAACAGCCCAACAAACAAACCCGCTGCA 471  
Qy 361 GCTCTGTTTATTAATGAGTATGCCGATGCCAGCTTACACACCTGGTGAAGAGATGCT 420  
Db 472 GCTCTGTTTATTAATGAGTATGCCGATGCCAGCTTACACACCTGGTGAAGAGATGCT 531  
Qy 421 CAAAGAACAGCCCTCTCAAGAAAAAGTTGGTAGAGGAGATCTCTCTCAACCCGAAAGCC 480  
Db 532 CAAAGAACAGCCCTCTCAAGAAAAAGTTGGTAGAGGAGATCTCTCTCAACCCGAAAGCC 591  
Qy 481 AGCCCAACAACTGCAAAAGCCCAAGCTGTACCAACAGTAAAGAAAGCGATGATAGGCA 540  
Db 592 AGCCCAACAACTGCAAAAGCCCAAGCTGTACCAACAGTAAAGAAAGCGATGATAGGCA 651  
Qy 541 ACAGAACATTAATCAGAGCTGTGTGGTTCAAGTCAAAAGATGCCCTTTAACAGAGTAC 600  
Db 652 ACAGAACATTAATCAGAGCTGTGTGGTTCAAGTCAAAAGATGCCCTTTAACAGAGTAC 711  
Qy 601 TTAAGCGAAATTAACCTTCCAAACAGCATAGATTCATACAGATCGACTCTATCTCCTG 660  
Db 712 TTAAGCGAAATTAACCTTCCAAACAGCATAGATTCATACAGATCGACTCTATCTCCTG 771  
Qy 661 TGGCTCTTGTGTCACCTCTTGGCCTTAACTGGAACTGCTGTTTATACCACTGCGCCTC 720  
Db 772 TGGCTCTTGTGTCACCTCTTGGCCTTAACTGGAACTGCTGTTTATACCACTGCGCCTC 831  
Qy 721 GTCCTCCCATATCAAAACCGCAGACACATACACTACTGGCTTATTCGGGACATCATATGT 780  
Db 832 GTCCTCCCATATCAAAACCGCAGACACATACACTACTGGCTTATTCGGGACATCATATGT 891  
Qy 781 GATATCATCTACCTTTATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAAAGAGCA 840  
Db 892 GATATCATCTACCTTTATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAAAGAGCA 951  
Qy 841 GGAGACATAATAGTGGATTCAAAATGAGCTAAGGAAACACTACAGGACTTCTTACAAAAATTT 900

952	Db	GGAGACATAATAGTGGATTCAAAATGAGCTAAGGAAACACTACAGGACTTCTACAAAATTT	1011
901	Qy	CAGTTGGATGTCGCATCAATAATAACCAATTTGATATTTGCTACCTCTTCTTTGGTTTAAT	960
1012	Db	CAGTTGGATGTCGCATCAATAATAACCAATTTGATATTTGCTACCTCTTCTTTGGTTTAAT	1071
961	Qy	CCAATGTTTAGAGCAATAGGATGTTAAAGTACACCTTCATTTTTGAAATTAATCATCAC	1020
1072	Db	CCAATGTTTAGAGCAATAGGATGTTAAAGTACACCTTCATTTTTGAAATTAATCATCAC	1131
1021	Qy	CTAGAGTCTATATGGACAAAGCANATATCTACAGAGTTATTCGAACAACCTGATACTTG	1080
1132	Db	CTAGAGTCTATATGGACAAAGCANATATCTACAGAGTTATTCGAACAACCTGATACTTG	1191
1081	Qy	CTGTTTTATCTGCACATTAATGCTGCTGTTTTATTACTTGGGCTTCAAACTATGAAGAAAT	1140
1192	Db	CTGTTTTATCTGCACATTAATGCTGCTGTTTTATTACTTGGGCTTCAAACTATGAAGAAAT	1251
1141	Qy	GGCACTACTAGATGGGTGTATGATGGGGAAGGAAACGAGTATCTGAGATGTTATTATGG	1200
1252	Db	GGCACTACTAGATGGGTGTATGATGGGGAAGGAAACGAGTATCTGAGATGTTATTATGG	1311
1201	Qy	GCAGTTCGAACCTTTAAATTAACAATTGGTGGCTTCCAGAACCAAACTTTATTGGAAT	1260
1312	Db	GCAGTTCGAACCTTTAAATTAACAATTGGTGGCTTCCAGAACCAAACTTTATTGGAAT	1371
1261	Qy	GTTTTTCAACCTCTTGAATTTTTTTTCTCGGAGTTTTTGTGTTCTCCAGTTAAATGGTCAG	1320
1372	Db	GTTTTTCAACCTCTTGAATTTTTTTTCTCGGAGTTTTTGTGTTCTCCAGTTAAATGGTCAG	1431
1321	Qy	ATGAGAGATGTGATTTGGAGCAGCTACAGCCAATCAGAACTACTTCCGCGCTCGCATGGAT	1380
1432	Db	ATGAGAGATGTGATTTGGAGCAGCTACAGCCAATCAGAACTACTTCCGCGCTCGCATGGAT	1491
1381	Qy	GACACCATTTGCCTACATGAAACAATTAATCTCAATTCCTAAACTGTGTGCAAAAGCGAGTTCGG	1440
1492	Db	GACACCATTTGCCTACATGAAACAATTAATCTCAATTCCTAAACTGTGTGCAAAAGCGAGTTCGG	1551
1441	Qy	ACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTAGATGAGTCTGATTTGCTTTAAG	1500
1552	Db	ACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTAGATGAGTCTGATTTGCTTTAAG	1611
1501	Qy	ACCTTACCAACTACGGTCCAGTTAGCCCTCGCCCATTTGATGTGAACCTTCAGCATCATCAGC	1560
1612	Db	ACCTTACCAACTACGGTCCAGTTAGCCCTCGCCCATTTGATGTGAACCTTCAGCATCATCAGC	1671
1561	Qy	AAAGTCGACTTGTTCAGGGTTGTGATACACAGATGATTTATGACATGTTGCTTAAGATTG	1620
1672	Db	AAAGTCGACTTGTTCAGGGTTGTGATACACAGATGATTTATGACATGTTGCTTAAGATTG	1731
1621	Qy	AAATCCGGTTCTTATTTGCCCTGGTCACTTGTCTGCAAAAGGAGGAAATTTGCAAGGAA	1680
1732	Db	AAATCCGGTTCTTATTTGCCCTGGTCACTTGTCTGCAAAAGGAGGAAATTTGCAAGGAA	1791
1681	Qy	ATGTATATCATCAAGCATGGAGAAGTCCAAGTTCTTGGAGGCCCTGATGTGTAAGAGTT	1740
1792	Db	ATGTATATCATCAAGCATGGAGAAGTCCAAGTTCTTGGAGGCCCTGATGTGTAAGAGTT	1851
1741	Qy	CTGTTTACTCTGAAAGCTGGGTGGGTGTTGGAGAAATTCAGCCCTTAGCAGCAGGAGGA	1800
1852	Db	CTGTTTACTCTGAAAGCTGGGTGGGTGTTGGAGAAATTCAGCCCTTAGCAGCAGGAGGA	1911
1801	Qy	GGAAACCGTCGAACTGCGCAATGTGGTGGCCCCACGGGTTGGCAATCTTTAACTCTAGAC	1860
1912	Db	GGAAACCGTCGAACTGCGCAATGTGGTGGCCCCACGGGTTGGCAATCTTTAACTCTAGAC	1971
1861	Qy	AAAAAGACCCCTCCAAAGAAATTTCTAGTGCATTTATCCAGATTTCTGAAAGGATCCTCATGAAG	1920
1972	Db	AAAAAGACCCCTCCAAAGAAATTTCTAGTGCATTTATCCAGATTTCTGAAAGGATCCTCATGAAG	2031
1921	Qy	AAAGCCAGAGTGCTTTTTTAAAGCAGAAGGCTTAAGACCGCAGAGCAACCCCTCCAGAAAA	1980

Db	2032	AAAGCCNAGAGTGCTTTTAAAGCAGAAAGGCTTAAGACCGCAGAACCAACCCCTCCAGAAAA	2098
Qy	1981	GATCTTGCCCTCCTCTTCCACCGAAAGAGACACCCAAACTGTTTAAAACTCTCCTTA	2040
Db	2092	GATCTTGCCCTCCTCTTCCACCGAAAGAGACACCCAAACTGTTTAAAACTCTCCTTA	2151
Qy	2041	GGAGGCACAGGAAAAAGCAAGTCTTTGCNAGACTACTCAAATTTGAAGCGAGAGCAAGCAGCT	2100
Db	2152	GGAGGCACAGGAAAAAGCAAGTCTTTGCNAGACTACTCAAATTTGAAGCGAGAGCAAGCAGCT	2211
Qy	2101	CAGAAGAAAGCAAAATTCCTGAAGGAGGAGAGGAAGAGGAAAAAGCAAAATGAAGATAAACAA	2160
Db	2212	CAGAAGAAAGCAAAATTCCTGAAGGAGGAGAGGAAGAGGAAAAAGCAAAATGAAGATAAACAA	2271
Qy	2161	AAAGAAAAATGAAGATAAAACAAAAAGAAAAATGAAGATAAAAGGAAAAAGAAAAATGAAGATAAA	2220
Db	2272	AAAGAAAAATGAAGATAAAACAAAAAGAAAAATGAAGATAAAAGGAAAAAGAAAAATGAAGATAAA	2331
Qy	2221	GATAAAGGAAGAGAGCCAGAGGAGAGAGAGCCACTGGACAGACTGAAATGTACAGCAAGTCCT	2280
Db	2332	GATAAAGGAAGAGAGCCAGAGGAGAGAGAGCCACTGGACAGACTGAAATGTACAGCAAGTCCT	2391
Qy	2281	ATTCCAGTGGAGGAAGAACCCCACTCAGTTAGAGGACAGCTTTTACCCAGAGGACTTCT	2340
Db	2392	ATTCCAGTGGAGGAAGAACCCCACTCAGTTAGAGGACAGCTTTTACCCAGAGGACTTCT	2451
Qy	2341	CGTCAATCACTCATTTATCAGCATGGCTCTTCTGCTCAGGCGGAGAGAGAGTTCCTTACT	2400
Db	2452	CGTCAATCACTCATTTATCAGCATGGCTCTTCTGCTCAGGCGGAGAGAGAGTTCCTTACT	2511
Qy	2401	ATTGAAAGTCAAAGAAAAAGGCTTAAGCAATAA	2430
Db	2512	ATTGAAAGTCAAAGAAAAAGGCTTAAGCAATAA	2541
RESULT 4			
US-11-060-756-2833			
; Sequence 2833, Application US/11060756			
; Publication No. US20050221354A1			
; GENERAL INFORMATION:			
; APPLICANT: Wyeth			
; APPLICANT: Mounts, William Martin			
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles o			
; TITLE OF INVENTION: Target Genes			
; FILE REFERENCE: AM101083 (031895-042000)			
; CURRENT APPLICATION NUMBER: US/11/060,756			
; CURRENT FILING DATE: 2005-02-18			
; NUMBER OF SEQ ID NOS: 303284			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 2833			
; LENGTH: 1400			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-11-060-756-2833			

	Query Match	54.2%	Score 1316;	DB 10;	Length 1400;
	Best Local Similarity	98.9%;	Prod. No. 3.8e-274;		
	Matches 1341;	Conservative	0;	Mismatches	15; Gaps
QY	1075	TACTTGCTGTTTATTCTGCACATAATAGCCCTGTCTTTATTACTGGGCTTCAAACATATGAA	1134		
DB	1	TACTTGCTGTTTATTCTGCACATAATAGCCCTGTCTTTATTACTGGGCTTCAAACATATGAA	60		
QY	1135	GGAAITGGCACTACTAGATGGGTGTATGATGGGAAGAAACGAGTATCTGAGATGTTAT	1194		
DB	61	GGAAITGGCACTACTAGATGGGTGTATGATGGGAAGAAACGAGTATCTGAGATGTTAT	120		
QY	1195	TATTGGGAGTTTCGAACTTTAAATACCAATGGTGGCCCTTCAGAAACACAAACCTTTATTT	1254		
DB	121	TATTGGGAGTTTCGAACTTTAAATACCAATGGTGGCCCTTCAGAAACACAAACCTTTATTT	180		
QY	1255	GAAATTTGTTTTTCAAACCTCTGAAATTTTTTTTTTCTGGAGTTTTTGTGTTCTCCAGTTTAAAT	1314		

Db	181	GAATTGTTTTTCAACTCTTGAAATTTTTTTCTGGAGTTTTTGTGTTCTCCAGTTTAATT	240
Qy	1315	GGTCAGATGAGAGATGTGATTTGGAGCAGCTACAGCCAATCAGAACTACTTTCGCGCCTGC	1374
Db	241	GGTCAGATGAGAGATGTGATTTGGAGCAGCTACAGCCAATCAGAACTACTTTCGCGCCTGC	300
Qy	1375	ATGGATGACACCAATTCGCTACATGTAACAATTAATCTCATTCCTAAACTTGTGTGCAAAAGCGA	1434
Db	301	ATGGATGACACCAATTCGCTACATGTAACAATTAATCTCATTCCTAAACTTGTGTGCAAAAGCGA	360
Qy	1435	GTTCGGACTTGGTATGAATATACATGGGACTCTCAAGAATCTAGATGAGTCTGATTTG	1494
Db	361	GTTCGGACTTGGTATGAATATACATGGGACTCTCAAGAATCTAGATGAGTCTGATTTG	420
Qy	1495	CTTAAGACCCCTACCAACTACGGTCCAGTTAGCCCTGCCATTGATGTGAACCTTCAGCATC	1554
Db	421	CTTAAGACCCCTACCAACTACGGTCCAGTTAGCCCTGCCATTGATGTGAACCTTCAGCATC	480
Qy	1555	ATCAGCAAAAGTCGACTTGTTCGAGGGTGTGATACACAGATGATTTATGACATGTTGCTA	1614
Db	481	ATCAGCAAAAGTCGACTTGTTCGAGGGTGTGATACACAGATGATTTATGACATGTTGCTA	540
Qy	1615	AGATTGAAATCCGTTCTCTATTGCTGCTGAGTCTTGTCTGCAAAAAGGAGAAATTGCG	1674
Db	541	AGATTGAAATCCGTTCTCTATTGCTGCTGAGTCTTGTCTGCAAAAAGGAGAAATTGCG	600
Qy	1675	AAGGAAATGTATATCATCAAGCATGGAGAGTCCAAGTTCTTGGAGGCCCTGATGGTACT	1734
Db	601	AAGGAAATGTATATCATCAAGCATGGAGAGTCCAAGTTCTTGGAGGCCCTGATGGTACT	660
Qy	1735	AAAGTTCTGGTTACTCTGTGAAGCTGGTGGTGGTTTGGAGAAATCAGCCCTCTAGCAGCA	1794
Db	661	AAAGTTCTGGTTACTCTGTGAAGCTGGTGGTGGTTTGGAGAAATCAGCCCTCTAGCAGCA	705
Qy	1795	GGAGGAGGAAACCGTCGAATCCCAATGTGTGGGCCACGGTTTGCCCAATCTTTTAACCT	1854
Db	706	GGAGGAGGAAACCGTCGAATCCCAATGTGTGGGCCACGGTTTGCCCAATCTTTTAACCT	765
Qy	1855	CTAGACAAAAGACCCCTCCAAGAAATTCCTAGTGCAATTCCTAGATTTCTGAAAGGATCCTC	1914
Db	766	CTAGACAAAAGACCCCTCCAAGAAATTCCTAGTGCAATTCCTAGATTTCTGAAAGGATCCTC	825
Qy	1915	ATGAAGAAAGCCAGAGTGTCTTTAAAGCAGAGGCTAAGACCGCAGAGCAACCCCTCCA	1974
Db	826	ATGAAGAAAGCCAGAGTGTCTTTAAAGCAGAGGCTAAGACCGCAGAGCAACCCCTCCA	885
Qy	1975	AGAAAAGATCTTGGCCCTCTCTTCCACCGAAGAGAGACACCCAAACTGTTTAAACT	2034
Db	886	AGAAAAGATCTTGGCCCTCTCTTCCACCGAAGAGAGAGACACCCAAACTGTTTAAACT	945
Qy	2035	CTCCTAGGAGGCACAGGAAAGCAAGTCTTGAAGACTACTCAAAATTTGAAGCGAGAGCAA	2094
Db	946	CTCCTAGGAGGCACAGGAAAGCAAGTCTTGAAGACTACTCAAAATTTGAAGCGAGAGCAA	1005
Qy	2095	GCAGCTCAGAAAGCAAGAAATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT	2154
Db	1006	GCAGCTCAGAAAGCAAGAAATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT	1065
Qy	2155	AAACAAAAGAAAATGAAGATAAAACAAAAGAAAATGAAGATAAAAGAAAAGAAAATGAA	2214
Db	1066	AAACAAAAGAAAATGAAGATAAAACAAAAGAAAATGAAGATAAAAGAAAAGAAAATGAA	1125
Qy	2215	GATAAGATTAAGAGGAGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT	2274
Db	1126	GATAAGATTAAGAGGAGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT	1185
Qy	2275	AGTCCTATTGTCAGTGGAG	2334
Db	1186	AGTCCTATTGTCAGTGGAG	1245
Qy	2335	ACTTCTCGTCAATCACTCATTTATCAGCATGGTCTCTTCTGCTGAGGGCGGAGAGAGGTT	2394
Db	1246	ACTTCTCGTCAATCACTCATTTATCAGCATGGTCTCTTCTGCTGAGGGCGGAGAGAGGTT	1305

Qy	2395	CTTACTATTGAAGTCAAAAGAAAAGGCTAAAGCAATAA	2430
Db	1306	CTTACTATTGAAGTCAAAAGAAAAGGCTAAAGCAATAA	1341

RESULT 5  
US-11-060-756-7105  
; Sequence 7105, Application US/11060756  
; Publication No. US20050221354A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William Martin  
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug  
; FILE OF INVENTION: Target Genes  
; FILE REFERENCE: AM101083 (031896-042000)  
; CURRENT APPLICATION NUMBER: US/11/060,756  
; CURRENT FILING DATE: 2005-02-18  
; NUMBER OF SEQ ID NOS: 303284  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7105  
; LENGTH: 1400  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-060-756-7105

Query Match 54.2%; Score 1316; DB 10; Length 1400;  
Best Local Similarity 98.9%; Pred. No. 3.8e-274;  
Matches 1341; Conservative 0; Mismatches 0; Indels 15; Gaps 1;

Qy	1075	TACTTGCTGTTTATTCTGCACATTAATGCCCTGTGTTTATTACTGGGCTTCAAACTATGAA	1114
Db	1	TACTTGCTGTTTATTCTGCACATTAATGCCCTGTGTTTATTACTGGGCTTCAAACTATGAA	60
Qy	1135	GGAATTTGCACACTACTAGATGGTGTATGATCGGGAAGAAACGAGTATCTGAGATGTTAT	1194
Db	61	GGAATTTGCACACTACTAGATGGTGTATGATCGGGAAGAAACGAGTATCTGAGATGTTAT	120
Qy	1195	TATTGGGCAGTTCGAACTTTAATTACCAATTTGGTGGCCTTCCAGAACCCAAACTTTATTT	1254
Db	121	TATTGGGCAGTTCGAACTTTAATTACCAATTTGGTGGCCTTCCAGAACCCAAACTTTATTT	180
Qy	1255	GAAATGTTTTTCAACTCTTGAAATTTTTTTTCTGGAGTTTTTGTGTTCTCCAGTTTAATT	1314
Db	181	GAAATGTTTTTCAACTCTTGAAATTTTTTTTCTGGAGTTTTTGTGTTCTCCAGTTTAATT	240
Qy	1315	GGTCAGATGAGAGATGTGATTTGGAGCAGCTACAGCCAATCAGAACTACTTCCGCGCCTGC	1374
Db	241	GGTCAGATGAGAGATGTGATTTGGAGCAGCTACAGCCAATCAGAACTACTTCCGCGCCTGC	300
Qy	1375	ATGGATGACACCAATTCGCTACATGAACAATTAATCTCCTAAACTTTGTGCAAAAGCGA	1434
Db	301	ATGGATGACACCAATTCGCTACATGAACAATTAATCTCCTAAACTTTGTGCAAAAGCGA	360
Qy	1435	GTTCGGACTTGGTATGAATATACATGGGACTCTCAAGAAATCTAGATGAGTCTGATTTG	1494
Db	361	GTTCGGACTTGGTATGAATATACATGGGACTCTCAAGAAATCTAGATGAGTCTGATTTG	420
Qy	1495	CTTAAGACCCCTACCAACTACGGTCCAGTTAGCCCTCGCCATTGATGTGAACCTTCAGCATC	1554
Db	421	CTTAAGACCCCTACCAACTACGGTCCAGTTAGCCCTCGCCATTGATGTGAACCTTCAGCATC	480
Qy	1555	ATCAGCAAAAGTCGACTTGTTCGAGGGTGTGATACACAGATGATTTATGACATGTTGCTA	1614
Db	481	ATCAGCAAAAGTCGACTTGTTCGAGGGTGTGATACACAGATGATTTATGACATGTTGCTA	1674
Qy	1615	AGATTGAAATCCGTTCTCTATTGCTGCTGAGTCTTGTCTGCAAAAAGGAGAAATTGCG	1674
Db	541	AGATTGAAATCCGTTCTCTATTGCTGCTGAGTCTTGTCTGCAAAAAGGAGAAATTGCG	600
Qy	1675	AAGGAAATGTATATCATCAAGCATGGAGAGTCCAAGTTCTTGGAGGCCCTGATGGTACT	1734
Db	601	AAGGAAATGTATATCATCAAGCATGGAGAGTCCAAGTTCTTGGAGGCCCTGATGGTACT	660



; TITLE OF INVENTION: NUCLEOTIDE GATED (CNG) CHANNEL IN RECOMBINANT HOST  
; TITLE OF INVENTION: CELLS AND USE THEREOF IN CELL BASED ASSAYS TO IDENTIFY  
; FILE OF INVENTION: SNELL MODULATORS  
; FILE REFERENCE: 078003-0291567  
; CURRENT APPLICATION NUMBER: US/10/189,507  
; CURRENT FILING DATE: 2003-02-12  
; PRIOR APPLICATION NUMBER: 60/303,140  
; PRIOR FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/337,154  
; PRIOR FILING DATE: 2001-12-10  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2607  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-189-507-3

Query Match 23.8%; Score 577.4; DB 6; Length 2607;  
Best Local Similarity 64.3%; Pred. No. 2.9e-114;  
Matches 866; Conservative 0; Mismatches 481; Indels 0; Gaps 0;

QY 617 TTCCAAACAGCATAGATTCCATACAGATCGACTCTATCTCTGTGGCTTTGCTGTGCA 676  
DB 782 TTCCCAAGAGCAATGACCGCTGACCACTGATGTATGCTTATGCTGTGCTTCTTCGTGG 841  
QY 677 CTCCTTGCCTTAACTGGAACTGCTGTGTTTATACCACTGCGCTCTGCTTCCCATATCAAA 736  
DB 842 TGAATGCGCTTGAATGGAATGTTGGCTGATTTCCCGTGGCTTCCCTACCAGA 901  
QY 737 CCGCAGACAAATACACTACTGCTTATTCGGGACATCATATGTGATATCATCTACTCTTT 796  
DB 902 CCCCAGACAAATCCACCACTGGCTGCTGATGGATTTACCTATGCGACCTCATCTACTCC 961  
QY 797 ATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAGAGGAGGAGACATAATAGTGG 856  
DB 962 TGGACATACCGTGTTCAGACAGCGCTGCAGTTTGTAGAGCGGGGACATCATTTACGG 1021  
QY 857 ATTCAAAATGAGCTAAGGAAACACTACAGGACTTCTACAAATTTTCAGTTGGATGCGCAT 916  
DB 1022 ACAAAAGGACATGCGAAATAACTACCTGAAGTCTCGCGCTTCAAGATGGACCTGCTCA 1081  
QY 917 CAATTAATACCATTTGATTTGCTACTCTTCTTTGGGTTTAAATCAATGTTTAGACAA 976  
DB 1082 GCCTCTGCGCTTGGATTTTCTCTATTTGAAAGTCCGTGTGAACCCCTCTCCGCTGC 1141  
QY 977 ATAGGATGTTAAAGTACACTTCAATTTTGAATTTAATCATCACCTAGAGTCTATAATGG 1036  
DB 1142 CCGCTGTTTAAAGTACATGGCTTCTTCGAGTTTACAGCGCCTGGAAATCACTCTCA 1201  
QY 1037 ACAAGCATATATCTACAGAGTATTTGGAACAACTGGATACCTTCTGTGTTTATTTCTGCA 1096  
DB 1202 GCAAGCCTACGTACAGGGTTCATCAGGACACAGCCTACCTTCTCTACAGCTGCAIT 1261  
QY 1097 TTAATGCTGCTTTTATTAAGTGGCTTCAAACTATGAAGAAATGGCACTACTAGATGG 1156  
DB 1262 TGAATTCCTGCTTTTATTAAGTGGCTTCAAGGCTTCAAGGCTTCAAGTGGCTTCAAGTGG 1321  
QY 1157 TGTATGATGGGGAAGGAAACGAGTATCTCAGATGTTTATTTATTTGGGAGTTTCGAATTTAA 1216  
DB 1322 TTTACAGTGGCTGGGAAACAGTTATTTGCTGTTTACTTCTTGTGTTGAAGACCTCA 1381  
QY 1217 TTAACATTTGGTGGCTTCCAGAACCAAACTTTTATTTGAAATTTGTTTTCAACTCTTGA 1276  
DB 1382 TCACCATGGGGGCTGCTGACCCCAAGACACTCTTTGAAATTTGTTTCCAGCTGCTGA 1441  
QY 1277 ATTTTCTGAGTTTTTGTGTTTCTCCAGTTTAAATGGTCAGATGAGAGATGATGG 1336  
DB 1442 ATTTATTCAGGGCGCTTTTGTCTTCTGTGATGATCGGACAGATGAGAGATGATGGTAG 1501  
QY 1337 GAGCAGCTACACCAATCAGAACTACTCTCCGCGCTGATGATGACACCATTTGCTCTACA 1396  
DB 1502 GGGCCGCCACCGGGGACAGACTACTACCGAGCTGCAATGAGACGACCGTGAAGTACA 1561

## RESULT 8

US-10-159-563-147  
; Sequence 147, Application US/10159563  
; Publication No. US20040009154A1  
; GENERAL INFORMATION:  
; APPLICANT: Khan, Javed  
; APPLICANT: Ringner, Markus  
; APPLICANT: Peterson, Carsten  
; APPLICANT: Meltzer, Paul  
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR  
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS  
; FILE REFERENCE: 11613.56US11  
; CURRENT APPLICATION NUMBER: US/10/159,563  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: US 10/133,937  
; PRIOR FILING DATE: 2002-04-25  
; NUMBER OF SEQ ID NOS: 444  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 147  
; LENGTH: 4382  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-159-563-147

Query Match 23.8%; Score 577.4; DB 6; Length 4382;  
Best Local Similarity 64.3%; Pred. No. 3.7e-114;  
Matches 866; Conservative 0; Mismatches 481; Indels 0; Gaps 0;

QY 617 TTCCAAACAGCATAGATTCCATACAGATCGACTCTATCTCTGTGGCTTCTTCTGTGCA 676  
DB 1983 TTCCCAAGCATTTGACCGCTGACCAACCTGATGTATGTTCTTATGGCTGTTCTTCGTGG 2042

```
QY 677 CTCTTGCCCTATAAAGTGGAGTGGTGTATATACCACTGCGCCTCGCTTCCCATATCAAA 736
Db 2043 TGATGGCTTGGAAATGGAACTGTGGCTGATTCCTCGTGGCTTCCCTACACAGA 2102
QY 737 CCGCAGACAACTACACTACTGGCTTATTTGGCGACATCATATGTGATATCATCTACTCTT 796
Db 2103 CCGCGACAACTACCACTGGCTGCTGATGATTTACCTATGCGACCTCATCTACTCTCC 2162
QY 797 ATGATATGCTATTTATTCAGCCAGACTCCAGTTCCTAAGAGAGAGAGACATAATAGTGG 856
Db 2163 TGGACATACCGTGTTCAGACACCGCTGCAGTTCCTGAGAGCGGGACATCATTTACGG 2222
QY 857 ATTCAAATGAGCTAAGAAACACTACAGGACTTCTACAAAATTTCAAGTTGGATGTGCGAT 916
Db 2223 ACAAAGAGACATGCGAAATAACTACTGAAGTCTCGCGCTTCAAGATGGACCTGCTCA 2282
QY 917 CAATAATACCATTTGATATTTGCTACTCTTCTTTGGGTTTAATCCAATGTTTAGACAA 976
Db 2283 GCCTCTCGCCTTGGATTTTCTCTATTTGAAAGTCGGTGTGAACCCCTCTCCGCGCTGC 2342
QY 977 ATAGGATGTTAAAGTACACTTCAATTTTGAATTAATCATCACCTAGAGTCTATATGG 1036
Db 2343 CCGCTGTTTAAAGTACATGGCCTTCTTCGAGTTTAAAGCCGCTGGATCCATCCTCA 2402
QY 1037 ACAAAGCATATATCTACAGAGTTTATTCGAAACAACTGGATACTTGGCTGTTTATTCGCACA 1096
Db 2403 GCAAAGCTACGTGTACAGGTCATCAGGACCAACAGCTACCTTCTCTACAGCCTGCAAT 2462
QY 1097 TTAATGCTGTGTTTATTTACTGGCTTCAAACTATGAAGAAATGGCACTACTAGATGG 1156
Db 2463 TGAATTCCTGTCTTATTTACTGGGCATCGGCCCTATCAGGCGCTCGGCTCCACTCACTGGG 2522
QY 1157 TGTATGATGGGAGGAAACGAGTATCTGAGATGTTTATTTATGGCAGTTCGAACTTTAA 1216
Db 2523 TTTACAGTGGCGTGGGAAACAGTTATATTCGCTGTACTTGTGCTGTGAAGACCTCA 2582
QY 1217 TTACCAATGGTGGCCTTCCAGAACCAAACTTTATTTGAAATGTTGTTTTCAACTCTGA 1276
Db 2583 TCACCATCGGGGGCTGCTGTGACCCCAAGACACTCTTTGAAATGTTCTTCCAGCTGCTGA 2642
QY 1277 ATTTTTCCTGAGTTTTTGTGTTCTCCAGTTTAAATGTTGCTCAGATGAGAGATGTGATG 1336
Db 2643 ATTTATTTACGGGGCTGTTTGTCTTCTGTGATGATCGACACAGATGAGAGATGTGGTAG 2702
QY 1337 GAGCAGCTACAGCAATCAAGAACTACTTCCGCGCTGCATGGATGACACCATTTGCCTACA 1396
Db 2703 GGCGCGCACCGGGGACAGCTTACTACCGAGCTGCATGGACAGCACGCTGAGTACA 2762
QY 1397 TGAACAAATTAATCAATTCCTAAACTTTGTGCAAAAGCGAGTTCCGACTTGGTATGAATATA 1456
Db 2763 TGAATTTCTAAGATCCCAAGTCCGTGAGAACCGCGTCAAGACCTGGTAGAGTACA 2822
QY 1457 CATGGGACTCTAAGAAATGCTAGATGAGTCTGATTTGCTTTAAGACCTTACCAACTACGG 1516
Db 2823 CCTGGCACTCGCAAGGATGCTGAGTGAATGAGTCAGAGCTGATGGTGCAGCTTCCAGACA 2882
QY 1517 TCCAGTTAGCCCTCGCCATTTGATGTGAATTTTACAGATCATCAGCAAAAGTGCATTTTCA 1576
Db 2883 TGGCGCTGGACCTCGCCATCGAATGAACTTACACATCTGTTAGCAAAAGTGCACCTCTTC 2942
QY 1577 AGGTTGTGTATACACAGATGATTTATGACATGTTTGTCTAAGATTTGAAATCCGTTCTCTATT 1636
Db 2943 AGGCTGTGACCGGCAGATGATCTTTGACATGCTGGAAGAGCTTCGCTCTGTTGTCTACC 3002
QY 1637 TGCCTGTGTACTTGTCTGMAAAGGAGAAATTTGCAAGGAAATGATATCATCAAGC 1696
Db 3003 TGCCCAACGATATGTTGCAAGAGGGGGAGATCGGCCGTGGAGATGATCATCATCCAGG 3062
QY 1697 ATGAGAGAGTCCAAGTTCTTGGAGGCCCTGATGTGATCTAAGATTTCTGTTACTCTGAAG 1756
Db 3063 CAGGGCAAGTGCAGGTTCTTGGCGCGCCCTGATGAGGAAATCTGTGCTGGTGCAGCTGAAG 3122
```

```
QY 1757 CTGGGTCGGTCTTTGGAGAAATCAGCCTTCTAGCAGCAGGAGGAGAAACCGTCGAACCTG 1816
Db 3123 CTGGAATCTGTGTTGGAGAAATAAGCTTGTGCGCTGTTGGGGCGGGAAACCGCGCACCG 3182
QY 1817 CCAATGTGTGGTGGCCACCGGTTTGGCAATCTTTTAACTCTAGACAAAAGAGACCTTCCAA 1876
Db 3183 CCAACGTGGTGGCGCAGCGGTTTACCACTCTTCTCATCTGATTAAGAAGGACCTGCAATG 3242
QY 1877 AAATTTAGTGCATTTATCCAGATTTCTGAAAGGATCTCTATGAAGAAAGCCAGAGTCTTT 1936
Db 3243 AGATTTTGGTGCATTTATCTGAGTCTCAGAAGTTTACTTCGGAAGAAAGCCAGGCGCATGC 3302
QY 1937 TAAAGCAGAGGCTAAGACCGCAGAG 1963
Db 3303 TGAGAAGCAACAATAAGCCCAAGGAGG 3329

RESULT 9
US-10-450-763-27462
; Sequence 27462, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 27462
; LENGTH: 3811
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2009) .. (3808)
; OTHER INFORMATION: 99% homologous to Homo sapiens cyclic nucleotide-gated cation
; OTHER INFORMATION: channel, accession number LI5296, Smith-Waterman Score=3168.
US-10-450-763-27462

Query Match 23.6%; Score 573.6; DB 9; Length 3811;
Best Local Similarity 64.6%; Pred. No. 2.3e-113;
Matches 855; Conservative 0; Mismatches 469; Indels 0; Gaps 0;

QY 640 ACAGATCGACTCTATCTCTGTGGCTTCTTGTGTCACTTGTGCTTATACCTGGAACCTGC 699
Db 2009 ACAGACTGATGTATGTCCTATGCTGCTGTTCTTCTGCTGATGCTTGGAAATGGAACTGT 2068
QY 700 TGGTTTATACCACTGCGCTCTGCTTCCCATATCAAAACCGCAGACACATACACTACTGG 759
Db 2069 TGGCTGATTTCCGCTGGGCTTCCCTTACAGACCCCGGACAAACATCCCACTGG 2128
QY 760 CTATTTCGGACATCATATGTGATATCATCTACTTTTATGATATGCTATTTATTCAGCCCC 819
Db 2129 CTGCTGATGATTTACCTATGCGACCTCATCTACTTCTCTGGACATCACCCTGTTCCAGACA 2188
QY 820 AGACTCCAGTTTGAAGAGGAGGAGACATAAATAGTGGATTCAAATGAGCTAAGGAAACAC 879
Db 2189 CGCTCGAGTTTGTGAGGCGGGGACATCATTCAGCAAAAAGGACATCGCAATAAAC 2248
QY 880 TACAGACTTCTACAAAAATTTAGTTGGATGTCGCATCAATAATACCATTTGATATTTC 939
Db 2249 TACCTGAAGTCTCGCCGCTTCAAGATGGACCTCTCAGCCCTCGCTTGGATTTTCTC 2308
QY 940 TACCTCTTCTTGGTTTAAATCCAATGTTTATAGAGCAAAATAGGATGTTTAAAGTACACTTCA 999
Db 2309 TATTTGAAGTCGGTGTGAACCCCTCTCCGCTCGCCCTGTTTAAAGTACATCGCC 2368
```

QY	1000	TTTTTGAATTTAATCATCACCTAGAGTCTATAATGGACAAAGCATATATCTACAGAGTT	1059
Db	2369		
QY	1060	ATTCGACAACTGGATCTGCTGTTTATTCGACATTAATAGCCCTGTTTATTACTGG	1119
Db	2429		
QY	1120	GCTTCAAACTATGAAGGAATTCGCACCTACTAGATGGTGTATGATGGGGAAGGAACGAG	1179
Db	2489		
QY	1180	TATCTGAGATGTTATTTATTTGGCAGTTGGAACCTTTTAATTAACCATTTGGTGGCCCTTCAGAA	1239
Db	2549		
QY	1240	CCACAACTTTATTTGAATTTGTTTCAACTCTTGAATTTTCTCGAGTTTGTG	1299
Db	2609		
QY	1300	TTCTCCAGTTTAATTTGGTCAGATGAGATGTGATTGGAGCAGCTACAGCCAAATCAGAAC	1359
Db	2669		
QY	1360	TACTTCCGCGCTGCATGATGACACCATTTGCCCTACATGAACAAATTTACTTCCATAA	1419
Db	2729		
QY	1420	CTTTGTCAAAAGCGAGTTCCGACTTGGTATGAATATACATGGACTCTCAAGAATGCTA	1479
Db	2789		
QY	1480	GATGAGTCTGATTTGCTTAAGACCCCTACCAACTACCGTTCAGTTAGCCCTCGCCATTGAT	1539
Db	2849		
QY	1540	GTGAATCTCAGCATCATCAGCAAGTCGACTTGTTCAGGGTTGTGATACACAGATGATT	1599
Db	2909		
QY	1600	TATGACATGTTGCTAAGATTGAAATCCGTTCTCTATTGGCTGGTGACTTTGTCTGCATA	1659
Db	2969		
QY	1660	AAGGAGAAATTTGGCAAGGAAATGTATATCATCAAGCATGGAGAGTCCAAAGTTCTTGA	1719
Db	3029		
QY	1720	GGCCCTGTATGGTACTAAAGTTCTGGTTACTCTGAAAGCTGGGTCCGGTGTTCGAGAAATC	1779
Db	3089		
QY	1780	AGCCTTCTAGCAGCAGGAGGAAACCGTCGAACTGCAATGTGGTGCCCAACGGGTTT	1839
Db	3149		
QY	1840	GCCAATCTTTTAACTCTAGACAAAAGACCCCTCCAAAGAAATTTCTAGTGCAATTCAGAT	1899
Db	3209		
QY	1900	TCTGAAAGGATCCTCATGAAGAAAGCCAGAGTGTCTTTTAAAGCAGAGGCTTAAGACCGCA	1959
Db	3269		
QY	1960	GAAG	1963
Db	3329	GAGG	3332

Publication No. US20050196754A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 790CIP3/US  
CURRENT APPLICATION NUMBER: US/10/450,763  
PRIOR FILING DATE: 2003-06-11  
PRIOR APPLICATION NUMBER: PCT/US01/08631  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 09/540,217  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 09/649,167  
PRIOR FILING DATE: 2000-08-23  
NUMBER OF SEQ ID NOS: 60736  
SOFTWARE: Custom  
SEQ ID NO 11960  
LENGTH: 2244  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SIMILAR  
LOCATION: (1301..1648)  
OTHER INFORMATION: 99% homologous to Homo sapiens cone photoreceptor cyclic:  
nucleotide-gated channel beta subunit, accession number  
OTHER INFORMATION: AF272900.Smith-Waterman Score=917.  
US-10-450-763-11960

Query Match		21.3%	Score 518;	DB 9;	Length 2244;
Best Local Similarity		94.7%	Pred: No. 1.9e-101;		
Matches 536;		Conservative 0;	Mismatches 30;	Indels 0;	Gaps 0;
QY	287	GAACAGTCCGACGAGCAGGAAGGAAATGGACCCCGGAAAGAGGTCCAAACAGCCCAAAA	141		
Db	83	GCAAACTATACAGGACAGAAAAACCAACCGCATGTTCTCACACACAGCCCAAAA	142		
QY	347	ACAAACCGCTGCGAGCTCTGTTTATAATGATGATGCGATGCCAGTGCAGCTACACAACCTGG	406		
Db	143	ACAAACCGCTGCGAGCTCTGTTTATAATGATGATGCGATGCCAGTGCAGCTACACAACCTGG	202		
QY	407	TGAAAAGAAATGCGTCAAAAGAACAGCCCTCTCAAGAAAAAGTTGGTAGAGGAGATCTCT	466		
Db	203	TGAAAAGAAATGCGTCAAAAGAACAGCCCTCTCAAGAAAAAGTTGGTAGAGGAGATCTCT	262		
QY	467	CCTACCCGAAAGCCAGCCCAAACTGCAAAAGCCCAACCGCTGTACCCAGTAAAAAGAAA	526		
Db	263	CCTACCCGAAAGCCAGCCCAAACTGCAAAAGCCCAACCGCTGTACCCAGTAAAAAGAAA	322		
QY	527	GCGATGATAAGCCCAACAGAACATTTACTACAGGCTGTTGGTTCAAAGTCAAAAAGATGC	586		
Db	323	GCGATGATAAGCCCAACAGAACATTTACTACAGGCTGTTGGTTCAAAGTCAAAAAGATGC	382		
QY	587	CTTTAAACAGAGTACTTTAAAGCGAATTTAAACTTCCAAACAGATAGATTCATACACAGATC	646		
Db	393	CTTTAAACAGAGTACTTTAAAGCGAATTTAAACTTCCAAACAGATAGATTCATACACAGATC	442		
QY	647	GACTCTATCTCTGTGGCTCTTGTGTTGTCACCTTTTGGCCCTATAACTGGAACCTGCTGTTTA	706		
Db	443	GACTCTATCTCTGTGGCTCTTGTGTTGTCACCTTTTGGCCCTATAACTGGAACCTGCTGTTTA	502		
QY	707	TACCACTGGGCTCGTCTTCCCATATCAAAACCGGAGACAAACATACACTACTGGCTTATTG	766		
Db	503	TACCACTGGGCTCGTCTTCCCATATCAAAACCGGAGACAAACATACACTACTGGCTTATTG	462		
QY	767	CGGACATCATATGTCATATCATCTTACCTTTATGATATCTATTATTCAGCTTCTTCT	826		
Db	563	CGGACATCATATGTCATATCATCTTACCTTTATGATATCTATTATTCAGCTTCTTCT	622		
QY	827	AGTTTGTAAAGGAGGAGACATAATA	852		
Db	623	AGTTTGTAAAGGAGGAGACATAATA	648		





```

Db      2212  TCGTTATCGGTGAGATCAGGATATCATCTCGACAGCCACTCGGATAAGCATGAGTATC 2271
Qy      1367  GGGCTGCGATGATGACACACATTTGGCTCATATGAAACAAATTTACTCCATTTCTTAAATTTGTC 1426
Db      2272  GCGAGCTGGAGGATGAGACACTGGAGTACATGCGAGCCCTCAATCTCTCGCAGGAGGTGC 2331
Qy      1427  AAAAGGAGTTCGGACTTGGTATGAATATACATGGGACTCTCAAGAGATGCTAGATGAT 1486
Db      2332  AGTCGAGGCTCAAAATGTTGGTTCCAGTTTACATGGGAGCAGCGCACTTTGGATGAAT 2391
Qy      1487  CTGATTTGCTTAAGACCTTACCACCTACGAGTCCAGTTCAGCTTCGCTTGGATGTAAGT 1546
Db      2392  CCAACATATGGATGCTTTGCCCATCAATTTGAAACCCGATATCGCCATCTTCGTCGACA 2451
Qy      1547  TCAGCATCATCAGCAAGTTCGACTTGTTCGAGGTTGTGATACACAGATGATTTATGACA 1606
Db      2452  TCCAAAGCTCTCCAAAGTGCACCTTTGCTGATTTGCGAGGAGGCTTTTACTCAGAGATC 2511
Qy      1607  TGTGCTAAGATGAAATCGGTTCTCTATTTGCTGCTGGTGAATTTGCTGCTGAAAGGGAG 1666
Db      2512  TGGTTCTAAACCTAAGGCGAGTCACATTTTTCGAGGAGATTTGCTTTGCGCAAGGCG 2571
Qy      1667  AAATGGCAAGAAATGATATCATCAAGCATGAGAGTCCAAAGTTCCTTGGAGGCCCTG 1726
Db      2572  AAGTGGCGAGAGAGATGATATCGTGAACCTGGGACAGGTTCAGGTGATGGGTGGTCCCTA 2631
Qy      1727  ATGGTACTAAAGTTCCTGGTTACTCTGAAAGCTGGGTGGTGGTGGAGAAATCAGGCTTC 1786
Db      2632  GTAGCATGATGCTGGCCACTCTAAACGAGGTTCCGTTTCGGGGAGATCAGTCTGC 2691
Qy      1787  TAGCAGCAGGAGAGAAACCGTCGAACCTGCAATGTGGTGGCCACGGGTTGGCAATC 1846
Db      2692  TGGGAATAAATGAGCGCGATCGCAGAACAGCAGATGTCCGATCCAAAGGCTACTCCAAAC 2751
Qy      1847  TTTTAACTCTAGCAAAAGACCCCTCCAGAAATTTCTAGTGCATTTACAGATTTCTGAAA 1906
Db      2752  TGTGCTGCTGCTCAATCGGATTTAAACGAGGTCAATGCTACTATCCACGCGCCAGG 2811
Qy      1907  GGATCCTCATGAAGAAAGCCAGAGTGTCTTTTAAAGCAGAGGC 1949
Db      2812  CGATTTCTCAAAAGCGTCTCGTCAGCTAATGCGMAAAATGC 2854

```

RESULT 15

```

US-10-172-118-427
; Sequence 427, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 427
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_000087
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-427

```

```

Query Match      7.3%; Score 176.4; DB 6; Length 2500;
Best Local Similarity 49.6%; Pred. No. 1.6e-27;
Matches 617; Conservative 0; Mismatches 591; Indels 36; Gaps 5;

740  CAGACAAACATACACTACTGGCTTATTTGGCGACATCATATGTGATATCATCTACCTTTATG 799
605  CTGATTACCTAGATATTTGGCTCAATTTGGATTACGTATCAGACATAGTCTATTTAATCG 664
800  ATATGCTATTTATCAGCCCGACACTCCAGTTTGTAAAGAGGAGGAGACATAATAGTGGATT 859
665  ATATG---TTTGTAGCAACAAGGACAGGTTACCTAGAACAAGGACTGCTGGTAAAGGAAG 721
860  CAAATGAGCTAAGGAAACACTACAGACTTCTACAAAATTTTCAGTTTGGATGTCGCAATCA 919
722  AACTTAACTCAATAAATAAATAATCAATCCAACTTGCNAATTTAACTTGATTTCTGTCTAC 781
920  TAATACCACTTTGATATTTTGTCTACCTCTTCTTTGGGTTTA---ATCCAAATGTTTTCAG 976
782  TGAATACCACTGATTTGCTGTATTTTAACTTAGGTGGAACATATCCAGAAAATTAGATTAA 841
977  ATAGGATGTTAAAGTACACTTCAATTTTGTGAATTTAATCATCACCTAGAGTCTATATGCG 1036
842  ACAGGTTGTTAGCGTTCTCTCGTATGTTTGAAGTTCTTCCAGAGAACAGAAACAAAGCAA 901
1037  ACAAAGCATATATACAGAGTTATTTCGAAACAACTGAGTATCTGCTGTTTATTTCTGCACA 1096
902  ACTATCCAAACATCTTCAGGATTTCCAACTTGTATGTATATCGTCATCATTTATCCACT 961
1097  TTAATGCCCTGTGTTTATTAATGCTGGCTTCAAACCTATGAGGAAATTTGGCACTACTAGAT 1156
962  GGAATGCAATGCTGTTCTACTCTATTTCTTAAAGCTATTTGGAATTTGGAATGATACATGG 1021
1157  TGTGTA-----TGATGGGGAAGAAACGAGTATCTGAGATGTT 1192
1022  TCTACCTCATATTAATGATCTCGAATTTGGCGCTTTGGCTAGAAAATACGTATACAGCC 1081
1193  ATTATTTGGCAGTTTCGAACTTTTAAATTTACCAATTTGGTGGCTTCCAGAACCAAACTTTAT 1252
1082  TTTACTGGTCTACACTGACTTTTGAATTTTGAATTTGCTTAAATTTGCTTACCATCG 1141
1253  TTGAAATTTGTTTTCAACTCTTGAATTTTCTGGAGTTTTGTCTTCCAGTTTAA 1312
1142  CTGAGTATGCTTTTGGTGGTGTGATTTCTTAAATTTGAGTGTAAATTTTGTCTACCATCG 1201
1313  TTGGTCAGATGAGAGATGATTTGGAGCAGCTACAGCCAAATCAGAACTACTTCCGGCCCT 1372
1202  TTGGTAAACATAGGTTCTATGATTTCCAAACATGAATGCAGCCAGAGCAGAAATTTCAAGCAA 1261
1373  GCATGGATGACACCATTTGCTTACATGAACAAATTTACTCCATTTCCCTAAACTTTGTCMAAAGC 1432
1262  GAATTTGATGCTATCAAGCAATATATGATTTTCGAAATTTGAAGCAAGATATGGAAAGA 1321
1433  GAGTTCCGGACTTGGTATGAATATATATGAGTGGACTCTCAAGAAATGCTAGATGAGTCTGATT 1492
1322  GGGTTATTAATGTTTGTGACTCTCTGTGACCAACAAACAAACAAACAGTTGATGAGAAAGAG 1381
1493  TGCTTAAGACCCCTACCAACTACCGTCCAGTTAGCCCTTCGCCATTTGATGTGAACCTTCAGCA 1552
1382  TCTTAAAGTATCTACCTGATATAAATAAGAGCAGAAATTTGCCATCAACGTTTCACTTAGACA 1441
1553  TCATCAGCAAGTCGACTTGTTCAGGGTTGTATACAGAGTATTTATGACATGTTGTC 1612
1442  CATTTAAAGAGGTAACGATTTTGTGATTTGTGAAGCTGCTGTTTGGTGGAGTTGGTCT 1501
1613  TAAGATTGAAATCCGTTCTCTATTTTGTCTGCTGCTGACTTTTGTCTGCAAAAGGAGAAATTC 1672
1502  TGAATTTGCAACCCCAAGTCTACAGTCTCGAGATATATATTTGCAAGAAAGGGGATATCG 1561
1673  GCAAGGAAATGTATATCATCAAGCATGGAGAAGTCCAAAGTTCTTTGGAGGCCCTTGATCGTA 1732
1562  GACGAGAGATGTACATTTATCAAGGAAGGCAAACTCGCTGTTGGTGGCAGATGATGGAGTCA 1621
1733  CTAAAGTTCTGGT-----TACTCTGAAAGCTGGGTGGTGGTGGTGGAGAAATC--AGCCTTC 1786

```

Db	1622	CTCAGTTTGTGGTATTGAGCGATGGCAGCACCTTCGGTGAGATCAGCATTTCTTAACATTA	1681
Qy	1787	TAGCAGCAGGAGGAGGAAACCGTCGAACTGCCAATGTGGTGCCCAACGGGTTTGCCAATC	1846
Db	1682	AAGGGAGCAAGCTGGCAATCGAAGNACGGCCCAATATTAAAGTATTGGCTACTCAGACC	1741
Qy	1847	TTTTAACTCTAGACAAAAAGACCTTCCAAAGAAATTCTAGTGCAATTATCCAGATTCTTGAAA	1906
Db	1742	TGTTCTGTCTCTCAAAAGATGACCTCATGGAAGCTCTAACTGAGTACCCAGATGCCAAAA	1801
Qy	1907	GGATCCTCATGAAGAAGCCAGAGTGCTTTTAAAGCAGAGGCT	1950
Db	1802	CTATGCTAGAGAAAAAGGGAAGCAAAATTTTAATGAAGATGGT	1845

Search completed: December 23, 2005, 23:51:25  
Job time : 1738 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2005, 17:34:16 ; Search time 312 Seconds  
(without alignments)  
4041.860 Million cell updates/sec

Title: US-09-855-828-3  
Perfect score: 2430  
Sequence: 1 atgtttaaatcgctgacaaa.....aagaaaggctaaagcaataa 2430

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 4168288 seqs, 259477437 residues

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.New.\*  
1: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	162.2	6.7	841	6	US-10-750-185-26200, A
2	126.2	5.2	1603	6	Sequence 26195, A
C 3	80.4	3.3	1687	6	Sequence 47568, A
C 4	72.2	3.0	380749	6	Sequence 13216, A
C 5	71.6	2.9	600	6	Sequence 20334, A
C 6	71.6	2.9	1442	6	Sequence 26196, A
C 7	67	2.8	118996	6	Sequence 84, Appl
C 8	65.2	2.7	184868	7	Sequence 88, Appl
C 9	64.4	2.7	191350	6	Sequence 4, Appl
C 10	64	2.6	201	6	Sequence 21234, A
C 11	63.4	2.6	173602	7	Sequence 25, Appl
C 12	62.8	2.6	201	6	Sequence 21194, A
C 13	62.8	2.6	201	6	Sequence 21195, A
C 14	62.8	2.6	201	6	Sequence 21351, A
C 15	62.8	2.6	28037	6	Sequence 13220, A
C 16	61.8	2.5	65723	6	Sequence 13200, A
C 17	61.6	2.5	901	6	Sequence 47572, A
C 18	61.6	2.5	150468	7	Sequence 56, Appl
C 19	61.6	2.5	193789	7	Sequence 55, Appl
C 20	61.4	2.5	2138	6	Sequence 60698, A
C 21	61.2	2.5	201	6	Sequence 21291, A
C 22	61.2	2.5	207908	6	Sequence 21, Appl
C 23	61.2	2.5	212805	7	Sequence 19, Appl

ALIGNMENTS

RESULT 1

US-10-750-185-26200  
; Sequence 26200, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MW1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26200  
; LENGTH: 841  
; TYPE: DNA  
; ORGANISM: Bovine 19866881614227  
US-10-750-185-26200

Query Match	6.7%	Score 162.2	DB 6	Length 841
Best Local Similarity	77.3%	Pred. No. 1.9e-25		
Matches 197	Conservative	0	Mismatches 58	Indels 0
Gaps	0			
Qy	2176	AAACAAAAGAAAATGAAGATAAAGGAAAAGAAAATGAAGATAAAGGATAAAGGATAAAGGAGAG	2235	
Db	395	AAAGACCAAAACCAGGAGAGAGGATAAAGGAAAAGAACTCTGAAGATAAAGGAGAGCA	454	
Qy	2236	CCAGAAGAGAGGACCTGGACAGACCTGAATGTACAGCAAGTCTCTATTGCGAGTGAGGAA	2295	
Db	455	ACAGCAGAGAAACCACTTGAACATCCAAATGTGAAACGAGTTCATTTACAGCGGAGGAA	514	
Qy	2296	GAACCCCACTCAGTTAGAAGCAGATTTTACCCAGAGGGAGTCTCTCTCAATCACTCATTT	2355	
Db	515	GCGCCCCAGGCAACTAGAAAGGCGGTTTTCCCGCAGAGGAACTACCCGTCACTCCTCATC	574	
Qy	2356	ATCAGCATGCTCTCTCTGCTGAGGCGGAGAGAGAGGTTCTTACTATTGAGTCAAGAA	2415	
Db	575	ATCAGCATGCTCTCTCTGCTGAGGCGGAGAGAGGTTCTGACGATCGAAGTCAAGAA	634	

```
QY 2416 AAGGCTAAGCAATAA 2430
Db 635 AAGGCTAAGCAATAA 649

RESULT 2
US-10-750-185-26195
; Sequence 26195, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 26195
; LENGTH: 1603
; TYPE: DNA
; ORGANISM: Bovine 19866880906828
US-10-750-185-26195

Query Match 5.2%; Score 126.2; DB 6; Length 1603;
Best Local Similarity 77.9%; Pred. No. 1.1e-17;
Matches 152; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1291 GTTTTGTGTTCTCCAGTTTAATTCGTACAGATGAGATGATGATGGAGCAGCTACAGCC 1350
Db 732 GGTGTGGAGTGATCATATATTTTGGTTTTAGATGCAAGAGCTAATTTGGGCGGCTACAGCC 791

QY 1351 AATCAGAACTACTTCGGCGCCTGCATGGATGACACCATTCCTACATGAACAATTAATCTCC 1410
Db 792 AATCAGAACACTTCGGCATCTGCATGATCATACATTCGCTATATGATTAATCTTCC 851

QY 1411 ATTCTAACTTGTCAAAAGCGAGTTCCGACTTGGTATGAATATACATGGGACTCTCAA 1470
Db 852 ATTCTAAGACTGTGCAGAAATCCGGTCCGGACTTCGTATGAATATACATGGGCGCTCTCAA 911

QY 1471 AGAATGCTAGATGAG 1485
Db 912 GGAATGCTAGGTAAG 926

RESULT 3
US-10-750-185-47568/c
; Sequence 47568, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47568

QY 2416 AAGGCTAAGCAATAA 2430
Db 635 AAGGCTAAGCAATAA 649

RESULT 4
US-10-995-561-13216/c
; Sequence 13216, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13216
; LENGTH: 380749
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(380749)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1,
US-10-995-561-13216

Query Match 3.0%; Score 72.2; DB 6; Length 380749;
Best Local Similarity 56.2%; Pred. No. 2.5e-05;
Matches 131; Conservative 2; Mismatches 100; Indels 0; Gaps 0;

QY 2088 AGAGCAAGCAGCTCAGAAAGAAAGAAATTTCTGAAGGAGCAGAGGAAAGGAAAGAAAGAA 2147
Db 65936 AGAGCAAGACTCTGTCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 65877

QY 2148 TGAAGATAAACAAAGAAAGAAATGAAGATAAACAAGAAAGAAATGAAGATAAAGGAAAGAA 2207
Db 65876 AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 65817

QY 2208 AATGTAAGATAAAGATAAAGAAAGAGAGCCAGAGAGAGCCACTGGAGACACCTTGAATG 2267
Db 65816 GAAGGAARGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAA 65757

QY 2268 TACAGCAAGTCTCTATTGTCAGTGGAGGAAGAACCCCTCAGTTAGAACGACAG 2320
Db 65756 GAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 65704

RESULT 5
US-10-750-185-20334/c
; Sequence 20334, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
```



```
QY 2187 AAATGAAGATAAGGA---AAGAAATGAAGATAAAGTAAGGAAGAGACCCAGAGA 2243
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55612 AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 55671
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2244 GAAGCACTGGACAGACCTGAATGTACACCAAGTCTATTGCAGTGGAGG 2293
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55672 AAAGAAAGAAAGAACTCAGTGAGTTACAGAAACACAAATCCACAGAGGG 55721
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-10-857-780-4/c
; Sequence 4, Application US/10857780
; Publication NO. US20050272043A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: BRAUN, ANDREAS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: RENELAND, RIKARD HENRY
; APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF BREAST CANCER AND TREATMENTS
; FILE REFERENCE: SEQ-4069-CP
; CURRENT APPLICATION NUMBER: US/10/857,780
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: 10/723,681
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/490,234
; PRIOR FILING DATE: 2003-07-24
; PRIOR APPLICATION NUMBER: 60/525,239
; PRIOR FILING DATE: 2003-11-25
; NUMBER OF SEQ ID NOS: 4962
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 191350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16914)..(16914)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (19266)..(19266)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (26334)..(26334)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (34457)..(34472)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (38627)..(38628)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40555)..(40555)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (57355)..(57355)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (61207)..(61207)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (64980)..(64980)
```

```
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (64987)..(64987)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (81452)..(81452)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (95893)..(95893)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103359)..(103359)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117565)..(117565)
; OTHER INFORMATION: n is a, c, g, or t
US-10-857-780-4

Query Match 2.7%; Score 64.4; DB 6; Length 191350;
Best Local Similarity 64.4%; Pred. NO. 0.00087;
Matches 112; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

QY 2096 CAGCTCAGAGAAGAAATAATTCTGAAGGAGGAGAGGAAGAGGAA-AAGAAAAATCAAGAT 2154
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150285 CAGCAAAAAAAGAGAGAGGAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 150226
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2155 AAACAAAAAGAAATGAAGATAACAAAAAGAAAAATGAAGATAAAGGAAAAAGAAAAATCAA 2214
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150225 GAAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150166
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2215 GATAAGATAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2268
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150165 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150112
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-10-995-561-21234/c
; Sequence 21234, Application US/10995561
; Publication NO. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21234
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-21234

Query Match 2.6%; Score 64; DB 6; Length 201;
Best Local Similarity 58.6%; Pred. NO. 8e-05;
Matches 109; Conservative 1; Mismatches 76; Indels 0; Gaps 0;

QY 2070 ACTACTCAAAATTGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2129
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 ACTTTGAAAAATAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2130 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2189
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 72
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2190 TGAAGATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2249
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21351
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-21351
```

```
Query Match      2.6%; Score 62.8; DB 6; Length 201;
Best Local Similarity 59.6%; Pred.No.0.00014;
Matches 106; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 2070 ACTACTCAAATTGAAGCGAGAGCAAGCAGCTCAGAGAAAGAAAAATTCCTGAAGGAGGAGA 2129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 ACTTTGMAAATAAAGAAAGACGAGACGACGACGAGAGAGATTAAGAGGAGAGGAAGAAGA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2130 GGAAGAGGAAAGAAAAATGAAGATAAACAAGAAAGAAATGAAGATAAACAAGAAAAA 2189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 119 GGAAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2190 TGAAGATAAGCAAGAAAGAAATGAAGATAAAGCAAGAGAGCCAGAGAGAGAGAG 2247
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2
```

```
RESULT 15
US-10-995-561-13220/c
; Sequence 13220, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13220
; LENGTH: 28037
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13220
```

```
Query Match      2.6%; Score 62.8; DB 6; Length 28037;
Best Local Similarity 53.5%; Pred.No.0.00092;
Matches 121; Conservative 4; Mismatches 101; Indels 0; Gaps 0;

QY 2070 ACTACTCAAATTGAAGCGAGAGCAAGCAGCTCAGAGAAAGAAAAATTCCTGAAGGAGGAGA 2129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9276 ACTTTGMAAATAAAGAAAGACGAGACGACGAGAGAGATTAAGAGGAGAGGAAGAAGA 9217
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2130 GGAAGAGGAAAGAAAAATGAAGATAAACAAGAAAGAAATGAAGATAAACAAGAAAAA 2189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9216 GGAAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9157
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2190 TGAAGATAAGCAAGAAAGAAATGAAGATAAAGCAAGAGAGCCAGAGAGAGAGAGAGCC 2249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9156 RGARGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9097
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2250 ACTGGACAGACCTGAATGTACAGCAAGTCCTATTTCAGTGGAGGAA 2295
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9096 GGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9051
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: December 23, 2005, 23:56:47  
Job time : 316 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2005, 16:49:01 ; Search time 11791 Seconds  
(without alignments)  
11714.832 Million cell updates/sec

Title: US-09-855-828-3  
Perfect score: 2430  
Sequence: 1 agtcttaaatcgtgacaaa.....aagaaaaggctgaagcaataa 2430

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sts.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_htg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2430	100.0	4369	8 AF272900	Homo sapi
2	2049.4	84.3	2135	8 AF228520	Homo sapi
3	1682.2	69.2	2826	4 AF490511	Canis fam
4	1362.4	56.1	4710	9 MMU243572	Mus muscu
5	853.4	35.1	1350	6 CQ716885	Sequence
6	577.4	23.8	3025	8 HUMCNGCCB	Homo sapien
7	577.4	23.8	3408	8 HUMCNGCCA	Homo sapien
8	577.4	23.8	4033	8 HSU58837	Homo sapien
9	577.4	23.8	4382	8 AF042498	Homo sapi
10	575.8	23.7	4320	6 CQ715550	Sequence
11	575.2	23.7	4763	9 BC045114	Mus muscu
12	572.2	23.5	3083	4 AF074012	Bos tauru
13	572.2	23.5	3090	4 AF074013	Bos tauru
14	572.2	23.5	3253	4 BTCNGACGN	B. taurus mR
15	572.2	23.5	3290	4 AF074014	B. taurus mR
16	572.2	23.5	4282	4 BTRPCNGCL	B. taurus mR
17	569.6	23.4	4238	9 RNCNG41	Rattus no
18	568	23.4	3236	6 CQ880122	Sequence

19	568	23.4	3236	9 RNCNG43	AJ000515 Rattus no
20	568	23.4	3328	9 AF068572	AF068572 Rattus no
21	330	13.6	205816	8 AC013751	AC013751 Homo sapi
22	237.4	9.8	2511	2 CEU73476	U73476 Caenorhabdi
23	223	9.2	3275	6 CQ598668	Sequence
24	223	9.2	3358	2 BT001439	BT001439 Drosophill
25	217.8	9.0	3099	9 BC046523	BC046523 Mus muscu
26	196.2	8.1	441	9 AF015728	AF015728 Rattus no
27	187.4	7.7	1709	2 AY060725	AY060725 Drosophill
28	181.2	7.5	2061	4 SSU85404	US8404 Sus scrofa
29	181.2	7.5	2487	6 CQ716337	CQ716337 Sequence
30	181.2	7.5	2857	8 S42457	S42457 CNGC-rod ph
31	176.8	7.3	2697	4 CFACGMP	X99914 C.familiari
32	176.8	7.3	2717	4 CFU83905	AX9305 Canis famli
33	176.4	7.3	2500	6 AX930698	AX930698 Sequence
34	176.4	7.3	2500	8 HUMCNGMP	M84741 Human CGMP-
35	173.6	7.1	2682	4 BTCGMPCH	X51604 B.taurus RN
36	169.8	7.0	1674	8 AF547222	AF547222 Homo sapi
37	169.8	7.0	181804	8 AC107068	AC107068 Homo sapi
38	166.6	6.9	178672	14 AC161820	AC161820 Pan trogl
39	166.6	6.9	219112	14 AC164974	AC164974 Pan trogl
40	166.4	6.8	2297	9 MMU19717	U19717 Mus musculu
41	164.8	6.8	2474	9 MUSCNGC	M84742 Mus musculu
42	162.4	6.7	6312	2 AF091302	AF091302 Limulus p
43	162.2	6.7	229715	14 AC161933	AC161933 Bos tauru
44	162	6.7	274861	14 AC154999	AC154999 Bos tauru
45	161.4	6.6	58409	9 AL772281	AL772281 Mouse DNA

#### ALIGNMENTS

RESULT 1	AF272900	4369 bp	mRNA	linear	PRI 29-AUG-2000
LOCUS	Homo sapiens cone photoreceptor cyclic nucleotide-gated channel				
DEFINITION	beta subunit (CNGB3) mRNA, complete cds.				
ACCESSION	AF272900				
VERSION	AF272900.1	GI:9247065			
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 (bases 1 to 4369)				
AUTHORS	Kohl,S., Baumann,B., Broghammer,M., Jagle,H., Sieving,P., Kellner,U., Spegal,R., Anastasi,M., Zrenner,E., Sharpe,L.T. and Wissinger,B.				
TITLE	Mutations in the CNGB3 gene encoding the beta-subunit of the cone photoreceptor cGMP-gated channel are responsible for achromatopsia (ACHM3) linked to chromosome 8q21				
JOURNAL	Hum. Mol. Genet. 9 (14), 2107-2116 (2000)				
PUBMED	10958649				
REFERENCE	2 (bases 1 to 4369)				
AUTHORS	Wissinger,B. and Kohl,S.				
JOURNAL	Direct Submission				
FEATURES	Submitted (30-MAY-2000) University Eye Hospital, Molecular Genetics Laboratory, Auf der Morgenstelle 15, Tuebingen D-72076, Germany				
source	Location/Qualifiers				
gene	1. .4369				
CDS	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	1. .4369				
	/gene="CNGB3"				
	47. .2476				
	/gene="CNGB3"				
	/codon_start=1				
	/product="cone photoreceptor cyclic nucleotide-gated channel beta subunit"				
	/protein_id="AAF86274.1"				
	/db_xref="GI:9247066"				

/translation="MFKSLTKVKNVKPIGNNENBQSSRRNEBESHPSNQSOQTAEQ  
ENKBEKLTSTVTEEPETHNQDKLSKNSSGDLTNPDPQNAEPTGTVEPEQ  
ENDPGKPNPSPAPVINEYADQHLNLRKMRQRTALYKXILVDSYDLSPEA  
SPQAKPTAVPVKESDDKPTBHYVRLMLFKVKMPLTEYLKRIKLPSINDSYDRLY  
LLMLLVTLAYNNWCFIPRLVFPVQADNHYWLIADICDIYLVDMFLPIORLQ  
FVRGDIIVDSNELRKHRTSKFQLDVNSIIPDICYLFFGFNFMFRANMLKTSF  
FEFNHLSIMDKATLIVIRITGYLLFILHINACVYVWASNYEGITGRWYDDEGN  
EYLRCYVAVARTLITIGLPEPOTLFEIVQLNFPSVFFSLIGQMRDVIAGATA  
NONYFRACWDDTIAMNNYSIPKLVQRVRYWETWDSQRMDSDDLKLTPTTVQL  
ALAI DVNFSIISKVDFKQDMOIMVLMRLKLSVLYLPDPFVCKKGEIKEMYI1KH  
GEVOVLGPDGTKVLTALKAGSVFGEISILAAAGGNRRRTANVAVGHPANLLTLDKXIL  
QELVHYDPDSERILMKKARVLKQAKTAETPPRDLALLFPFKBEETPKLTKLUGG  
TKGASLRLKLRQQAQKNSGEBEKENEDKOKENEDKOKENEDKOKENEDK  
DKGREPEKPLDRPECTASPIAVEBEPHSVRVTLPRGTSRQSLIISMAPSAGEEYV  
LTIEVKEKAKQ"

ORIGIN

Query Match		100.0%;	Score 2430;	DB 8;	Length 4369;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2430;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGTTTAAATCGCTGACAAAGTCAACAAGGTGAAGCCTATAGGAGAGAACAAATGAGAAT	60		
DB	47	ATGTTTAAATCGCTGACAAAGTCAACAAGGTGAAGCCTATAGGAGAGAACAAATGAGAAT	106		
QY	61	GAACAAAGTTCGTGCGAATGAAGAAGCTCTCACCAAGTAATCAGTCTCAGCAAAACC	120		
DB	107	GAACAAAGTTCGTGCGAATGAAGAAGCTCTCACCAAGTAATCAGTCTCAGCAAAACC	166		
QY	121	ACAGCACAGGAGAAAAACAAGGTGAAGAGAAATCTCTCAAAACCNAAGTCAACTCAGTC	180		
DB	167	ACAGCACAGGAGAAAAACAAGGTGAAGAGAAATCTCTCAAAACCNAAGTCAACTCAGTC	226		
QY	181	ACGTCCTGAAGAGCCACACCAACATACAAAGCAAACTCTCCAAGAAAAATTCCTCTGGA	240		
DB	227	ACGTCCTGAAGAGCCACACCAACATACAAAGCAAACTCTCCAAGAAAAATTCCTCTGGA	286		
QY	241	GATCTGACCAAAACCCGTGACCCCTCAAAATGAGCAGCAACCAACTGGAACTGGCAGAG	300		
DB	287	GATCTGACCAAAACCCGTGACCCCTCAAAATGAGCAGCAACCAACTGGAACTGGCAGAG	346		
QY	301	CAGAAGGAATGGAACCCCGGGAAGAAGGTCAAAAGCCCAACAAACCAACCCGCTGGA	360		
DB	347	CAGAAGGAATGGAACCCCGGGAAGAAGGTCAAAAGCCCAACAAACCAACCCGCTGGA	406		
QY	361	GCTCTCTGTTATAAATGAGTATGCCGATGCCAGCTACACAACTGGTGAAGAAGATGGCT	420		
DB	407	GCTCTCTGTTATAAATGAGTATGCCGATGCCAGCTACACAACTGGTGAAGAAGATGGCT	466		
QY	421	CAAGAAGACGCCCTCTTACAAGAAAAAGTTGGTAGAGGGAGATCTCTCTCAACCCGAAAGCC	480		
DB	467	CAAGAAGACGCCCTCTTACAAGAAAAAGTTGGTAGAGGGAGATCTCTCTCAACCCGAAAGCC	526		
QY	481	AGCCCAACAACCTGCAAGGCCACCGCTGTACACCGTAAAGAAAGCGATGATAGCCA	540		
DB	527	AGCCCAACAACCTGCAAGGCCACCGCTGTACACCGTAAAGAAAGCGATGATAGCCA	586		
QY	541	ACAGAACAATTACTACAGGCTGTGTGGTTCAAAAGTCAAAAGATGCTTTTAAACAGAGTAC	600		
DB	587	ACAGAACAATTACTACAGGCTGTGTGGTTCAAAAGTCAAAAGATGCTTTTAAACAGAGTAC	646		
QY	601	TTAAAGCGAATTAACCTTCAAAACAGCATAGATTTCATACACAGATCGACTCTATCTCCTG	660		
DB	647	TTAAAGCGAATTAACCTTCAAAACAGCATAGATTTCATACACAGATCGACTCTATCTCCTG	706		
QY	661	TGGCTCTTGCTTGTGACTCTTGCTTAACTGGAAGTGTGTTTATACCACTGGCCCTC	720		
DB	707	TGGCTCTTGCTTGTGACTCTTGCTTAACTGGAAGTGTGTTTATACCACTGGCCCTC	766		
QY	721	GTCTTCCCATATCAAAACCGCAGACAACTACTACTGCTTATTCGGGACATCATATGT	780		
DB	767	GTCTTCCCATATCAAAACCGCAGACAACTACTACTGCTTATTCGGGACATCATATGT	826		

QY	781	GATATCATCTACCTTTATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAAAGAGA	840
DB	827	GATATCATCTACCTTTATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAAAGAGA	886
QY	841	GGAGACATAANTAGTGGATTCAATAGAGCTAAGGAACAACATACAGGACTTCTACAAAAATTT	900
DB	887	GGAGACATAANTAGTGGATTCAATAGAGCTAAGGAACAACATACAGGACTTCTACAAAAATTT	946
QY	901	CAGTTGGATGTCGCATCAATAATACCAATTCATATTTCTCTACCTCTCTCTTTGGGTTTAAAT	960
DB	947	CAGTTGGATGTCGCATCAATAATACCAATTCATATTTCTCTACCTCTCTCTTTGGGTTTAAAT	1006
QY	961	CCAAATGTTTAGAGCAAAATAGGATGTTAAAGTACACTTTCATTTTTTGAATTTAAATCATCAC	1020
DB	1007	CCAATGTTTAGAGCAAAATAGGATGTTAAAGTACACTTTCATTTTTTGAATTTAAATCATCAC	1066
QY	1021	CTAGAGTCTATAATGGCAAAAGCATATATCTACAGAGTTATTCGAACAACCTGGATCTTG	1080
DB	1067	CTAGAGTCTATAATGGCAAAAGCATATATCTACAGAGTTATTCGAACAACCTGGATCTTG	1126
QY	1081	CTGTTTATTTCTGCACATTAATGCCTGTGTTTATTACTGGGCTTCAAACTATGAAGGAATTT	1140
DB	1127	CTGTTTATTTCTGCACATTAATGCCTGTGTTTATTACTGGGCTTCAAACTATGAAGGAATTT	1186
QY	1141	GGCACTACTAGATGGGTCTATGATGGGAAGGAAACGAGTATCTGAGATGTTTATTATTGG	1200
DB	1187	GGCACTACTAGATGGGTCTATGATGGGAAGGAAACGAGTATCTGAGATGTTTATTATTGG	1246
QY	1201	GCAGTTCGAACTTTAATTTACCATTGGTGGCCCTTCAGAAACCAACAACTTTATTGAAATTT	1260
DB	1247	GCAGTTCGAACTTTAATTTACCATTGGTGGCCCTTCAGAAACCAACAACTTTATTGAAATTT	1306
QY	1261	GTTTTTCAACTCTTGAAATTTTTTTCTGGAGTTTTTGTGTTCTCCAGTTTAAATTCGTCAAG	1320
DB	1307	GTTTTTCAACTCTTGAAATTTTTTTCTGGAGTTTTTGTGTTCTCCAGTTTAAATTCGTCAAG	1366
QY	1321	ATGAGAGATGTAATGGAGCAGCTACAGCCAATCAGAACTACTTCCGCGCTGCAATGAT	1380
DB	1367	ATGAGAGATGTAATGGAGCAGCTACAGCCAATCAGAACTACTTCCGCGCTGCAATGAT	1426
QY	1381	GACACCATTCCTACATGAACTACTCTCAATCTCTAACTTGTGCAAAAGCGAGTTCTGG	1440
DB	1427	GACACCATTCCTACATGAACTACTCTCAATCTCTAACTTGTGCAAAAGCGAGTTCTGG	1486
QY	1441	ACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTAGATGAGTCTGATTGCTTAAAG	1500
DB	1487	ACTTGGTATGAATATACATGGGACTCTCAAGAAATGCTAGATGAGTCTGATTGCTTAAAG	1546
QY	1501	ACCCTACCAACTACGGTCCAGTTAGCCCTCGCCATTGATGTGAACCTTCAGCATCATCAGC	1560
DB	1547	ACCCTACCAACTACGGTCCAGTTAGCCCTCGCCATTGATGTGAACCTTCAGCATCATCAGC	1606
QY	1561	AAAGTCGACTTGTTCAGGGTTGTGATACAGATGATTTATGACATGTTGCTAAGATTG	1620
DB	1607	AAAGTCGACTTGTTCAGGGTTGTGATACAGATGATTTATGACATGTTGCTAAGATTG	1666
QY	1621	AAATCCGTTCTCTATTTTGCCTGGTGACTTGTCTGCAAAAAGGGAGAAATTTGGCAAGAA	1680
DB	1667	AAATCCGTTCTCTATTTTGCCTGGTGACTTGTCTGCAAAAAGGGAGAAATTTGGCAAGAA	1726
QY	1681	ATGATATCATCAAGCATGGAGAAAGTCCAAGTTCTTGGAGCCCTTGATGGTACTTAAAGTT	1740
DB	1727	ATGATATCATCAAGCATGGAGAAAGTCCAAGTTCTTGGAGCCCTTGATGGTACTTAAAGTT	1786
QY	1741	CTGGTACTCTGAAAGCTGGGTGGTGTTCGAGAAATTCAGCCTTCTAGCAGCAGGAGGA	1800
DB	1787	CTGGTACTCTGAAAGCTGGGTGGTGTTCGAGAAATTCAGCCTTCTAGCAGCAGGAGGA	1846
QY	1801	GGAAACCGTCCAACTGCCAATGTGGTGGCCACAGGGTTTGCCAACTCTTTAACTCTAGAC	1860
DB	1847	GGAAACCGTCCAACTGCCAATGTGGTGGCCACAGGGTTTGCCAACTCTTTAACTCTAGAC	1906
QY	1861	AAAAAGACCCCTCCAAGAAATTTCTAGTGCATTTATCCAGATTCTGAAGAGGATCCTCATGAAG	1920





HYPDSEKLMKXSVLLKKKAPATETTPPRKGLAFLFPPKQETPKIFKALLGGTCKAG  
LTRLKLRKEQITQKTSNREGGKREYEDKREPSKILDSSECRANCIIBEMP  
QSTRRAALPRGTTTQSLIISMSPSAEAGEVLITIEVKEKAKQ

## ORIGIN

Query Match	69.2%;	Score	1682.2;	DB	4;	Length	2826;	
Best Local Similarity	83.8%;	Pred. No.	2.4e-286;					
Matches 1972;	Conservative	0;	Mismatches	312;	Indels	75;	Gaps	3;

  

QY	78	GAATGAAGAGGCTCTCACCAAGTAATCAGTCTCAGCAAAACACACAGCACAGGAGAGAAAA	137
DB	84	GAATAAACAGATCCTGATCCAGCAATCAGCTCAGCAATCTACAAGACAGGAGAAAA	143
QY	138	CAAAGGTGAAGAGAAATCTCTCAAAACCAAGTCAATCCAGTCACTCTGAAGAGCCACA	197
DB	144	CAAAAGTGAATAATAATCTCTCAAAACCAAGATGACTCCAGTCACTGTTGAAGAAATCACA	203
QY	198	CACCAACATACAGACAAACTCTCCAGAAATAATCCTCTGGAGATCTGACCAACCAACC	257
DB	204	TGCCAAATGCAAGATATAATCTCCAGAGAAAAATCCTCAGAGACCTGACGCAAAATCC	263
QY	258	TGACCTCTCAAAATGACAGCAGAACCAACTTGAACAGTGCCAGACAGAGAGAAATGGACCC	317
DB	264	CAACCATCAACATCCAACAGATCAAGAGGGGCAATGTCAAGCAGAGAGAAATGGAAAC	323
QY	318	CGGGAAGAGAGTCCAAACAGCCCAACAAAACCAACCCCTGCGAGCTCTCTGTTATAAATGA	377
DB	324	TGGGAAGAGAGGCTAGTCAGGCCCAAAAGCAAAACCCCTTGGAGTCCCTGTTATAAATGA	383
QY	378	GTATGCCGATGCCAGTACACAACTGTTGTAAGAAATGCGTCAAGAGACAGCCCTCTTA	437
DB	384	GTATGCTGATGCTCAGCTACACAACTGTTGTAAGAAATGCGTCAAGAGACAAATGCTCTA	443
QY	438	CAAGAAAAAGTGTGAGAGGAGATCTCTCTCACCAGAGCCAGCCCAACAACTGCAAA	497
DB	444	TAAGAAAAAGTGTGAGAGGAGATATATCTCCTCAGTGAAGCTAGACCCCAAACTGCAAA	503
QY	498	GCCACGCGCTGACCAACAGTAAAGAAAGCGATGATAGCCAA---CAGAACTACTTA	554
DB	504	GCCACAGCGCTACCATCAACACAGAGAGCAATGCTAAGCTAAAGAAAGCACTACTTA	563
QY	555	CAGCTGTTGTTGTTTCAAGTCAAAAGATGCTCTTTAAACAGAGTACTTAAAGCGAATTA	614
DB	564	TCACATATTGTTGTTTAAATTCAGAGAGATGCTCTGACAGAGTACCTTAAACGATTA	623
QY	615	ACTTCCAAACAGCATAGATTATACACAGATCGACTCTATCTCTGTGGCTCTTGTCTGT	674
DB	624	ACTTCCAGGAGCATAGATTATACACAGATCGACTCTATCTCTGTGGCTCTTGTCTGT	683
QY	675	CACCTTTGCCCTATAACTGGAACTGCTGGTTTATACCACTCGCCTCGTCTTCCCATATCA	734
DB	684	CACCATTTGCCCTATAACTGGAACTGCTGGTTTATACCACTACGCTCGTCTTCCCATATCA	743
QY	735	AACGGCAGACAACTACACTGCTGCTTATGCGGACATCATATGATATCATCTACCT	794
DB	744	AACACAGACAACTACACTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGAT	803
QY	795	TTATGATATGCTATTTATCCAGCCAGACTCCAGTGTGTAAGAGAGGAGACATAATAGT	854
DB	804	TTGTGATATGCTATTTATCCAGCCAGACTCCAGTGTGTAAGAGAGGAGACATAATAGT	863
QY	855	GGATTCAAAATGAGCTAAGGAAACACTACAGACTCTTCAAAAATTTTCAGTTGGATGCGC	914
DB	864	GGATTCAAAATGAGCTAAGGAAACACTACAGACTCTTCAAAAATTTTCAGTTGGATGCGC	923
QY	915	ATCAATTAATCAATTTGATATTTGCTACCTCTTCTTGGGTTTAAATCCAATGTTTAGAGC	974
DB	924	GTCAGTAATGCCATTTGATGTTTTTTTACCTCTTCTTGGGTTTAAATCCAATGTTTAGAGT	983
QY	975	AAATAGAGTGTAAAGTACACTTCAATTTTTTGAATTTAAATCATCACCTAGAGTCTATAAT	1034
DB	984	GAATAGGATATTGAAGTACACTTCAATTTTTTGAATTTAAATCATCACCTAGAGTCTATAAT	1043

QY	1035	GGACAAAGCATATATCTACAGAGTTATTTCGACAACTCGATACTTGGTGTATTATCTGCA	1094
DB	1044	GGACAAAGCATATATCTACAGAGTTATTTCGACAACTCGATACTTGGTGTATTATCTGCA	1103
QY	1095	CATTAAATGCCTGTGTTTATTACTGGGCTTCAAATCTATGAAGGAATTTGGCACTACTAGATG	1154
DB	1104	CATTAAATGCCTGTGTTTATTACTGGGCTTCAAATCTATGAAGGAATTTGGCACTACTAGATG	1163
QY	1155	GCTGATGATGGGAGGAAACAGATATCTGAGATGTTTATTATGGGCACTTGGCACTTT	1214
DB	1164	GCTGATGATGGGAGGAAACAGATATCTGAGATGTTTATTATGGGCACTTGGCACTTT	1223
QY	1215	AATTACCATTTGCTGGCTTCCAGAACCAACACTTTATTGAAATTTGTTTTTCAACTCTT	1274
DB	1224	AATTACCATTTGCTGGGCTTCCAGAACCAACACTTTATTGAAATTTGTTTTTCAACTCTT	1283
QY	1275	GAATTTTTTTCTGGAGTTTTTGTCTTCCAGTTTAAATTTGGTCAGATGAGAGATGTGAT	1334
DB	1284	GAATTTTTTTCTGGAGTTTTTGTCTTCCAGTTTAAATTTGGTCAGATGCAAGATCTAAT	1343
QY	1335	TGAGCAGCTACAGCCAAATCAGAACTACTTTCCGCGCTCGATGGATGACACCAATTCGCTA	1394
DB	1344	TGGGCGAGCCACAGCCAAATCAGAACTACTTTCCGCGCTCGATGGATGACACCAATTCGCTA	1403
QY	1395	CATGAACAATTAATCCATTTCTTAACTTCTGCAAAAGCGAGTTCCGACTTGGTATCAATA	1454
DB	1404	CATGAACAATTAATCCATTTCTTAACTTCTGCAAAAGCGAGTTCCGACTTGGTATCAATA	1463
QY	1455	TACATGGGACTCTCAAAGAAATGCTAGATGAGTCTGATTTGCTTAAAGACCTACCAACTAC	1514
DB	1464	TACATGGGACTCTCAAAGAAATGCTAGATGAGTCTGATCTGCTCGACCCCTGCTGCTAC	1523
QY	1515	GGTCCAGTTAGCCCTCGCCATTTGATGTAATTTAGCATCATCAGCAAAAGTTCGACTTGT	1574
DB	1524	CATGCAGTTAGCCCTCACTGTGATGTGAACCTCAGCATCATCAGCAAAAGTTCGACTTGT	1583
QY	1575	CAAGGTTGTGATACACAGATGATTTATGACATGTTGCTTAAAGTTCGACTTCTCTA	1634
DB	1584	CAAGGTTGTGATACACAGATGATTTATGACATGTTGCTTAAAGTTCGACTTCTCTA	1643
QY	1635	TTTGCCTGTGACTTTGCTGCAAAAGGAGAAATTTGGCAAGAAATGTATATCATCA	1694
DB	1644	TTTGCCTGTGACTTTGCTGCAAAAGGAGAAATTTGGCAAGAAATGTATATCATCA	1703
QY	1695	GCAATGAGAGTCCAAAGTTCTTGGAGCCCTGATGTAATAAGTTCTGGTTACTCTGAA	1754
DB	1704	GCAAGGAGAGTCCAAAGTTCTTGGAGCTTCTGATGCTGCCAAAGTTCTGGTTACTCTGAA	1763
QY	1755	AGCTGGGTGGGTTTGGGAAATCAGCTTCTAGCAGCAGGAGGAGGAAACCTCGAAC	1814
DB	1764	AGCTGGGAGAGTGTGTTGGGAAATCAGCTTCTAGCAGGAGGAGGAGGAAATCGCCGAC	1823
QY	1815	TGCCAAATGTTGGTGGCCACCGGTTTGCCAAATCTTTTAACTCTAGACAAAAAGACCTCCA	1874
DB	1824	TGCCAAATGTTGATAGCCATGGGTTTGCCAAATCTTTTAACTCTAGACAAAAAGACCTCCA	1883
QY	1875	AGAAATTTAGTCATTTATCCAGATTTCTGAAAAGCTCTCATGAAAAGCCAGTGTCT	1934
DB	1884	AGAAATTTAGTCATTTATCCAGATTTCTGAAAAGCTCTCATGAAAAGCCAGTGTCT	1943
QY	1935	TTTAAAGCAGAGGCTTAAGCCCGCAGAACCCCTCCAGAAAAAGATCTTCCCTCTCT	1994
DB	1944	TCTAAAGAGAGGCTTCGCGCCACAGACAACTCTCCAGAAAAAGGACTTGGCTTTCT	2003
QY	1995	CTTCCCAACCAAGAGAGACACCCCAACTGTTTAAAGCTCTCTAGGAGGACAGGAAA	2054
DB	2004	CTTCCCAACCAAGAGAGACACCCCAACTGTTTAAAGCTCTCTAGGAGGACAGGAAA	2063
QY	2055	AGCAAGTCTTTCAGAGTACTCAAAATTTGAACGAGAGCAAGCAGCTCAGAGAAA---GA	2111
DB	2064	AGCAGGACTTTCAGAGTACTCAAACTGAAGAGAGAAACAACTTTCAGAGAAAACAGTGA	2123
QY	2112	AAATTTCTGAAGGAGGAGAGAGGAAAGAAATGAAGAAATGAACAAAAAGAAATGA	2171



```
Db 914 AGATTCAAATGAGCTGAAGAGGAACCTACCGAGTCTCTCAAAAGTTTCGAGTGAATGGC 973
Qy 915 ATCAATAATACCAATTTGATATTTGCTACCTCTCTTTGGGTTTAAATCCAAATGTTAGAGC 974
Db 974 ATCTCTACTGCCAATTCGAAGTTCTCTACATCTCTTTGGAGTTAAATCCAAATATTCAGAGC 1033
Qy 975 AAATGAGATGTTAAAGTACACTTTCACTTTTGAATTTAATCATCATCAGTCTAGATCTATAAT 1034
Db 1034 AAATCGGATATTAAGATATATCTTCACTTTTGAAGTTTAAACCATCACTCGAGTCTATAAT 1093
Qy 1035 GGACAAAGCATATATCTACAGAGTATTCGAACAACTGGATATCTGCTGTTTATCTCGCA 1094
Db 1094 GGACAAAGCATATGCTACAGAGTCACTCGAACACTGGCTACTGCTGTTCTCTCTGCA 1153
Qy 1095 CATTAAATGCTGTGTTTATTACTGGGCTTCAAACTATGAAGGAATTTGGCACTACTAGATG 1154
Db 1154 CATTAAAGCTGTGTTTATTACTGGGCTTCAGACTATGAAGGAATTTGGCTCAACTAAATG 1213
Qy 1155 GGTGATGATGGGGAAGGAACGAGTATCTGAGATGTTATTATTTGGGAGTTCGAACTTT 1214
Db 1214 GGTCTATAATGGTGAAGGCAACAGTATCTGCGATGCTTTTATTTGGGCAAGTTTCAACTTT 1273
Qy 1215 AATTACCATTTGGTGGGCTTCCAGAACACAACTTTTATTTGAAATTTGTTTTCAACTCTT 1274
Db 1274 AATTACTATCGGGGGCTTCCAGAGCCACAGACTTCAITTTGAAATTTGTTTTCAAATCTT 1333
Qy 1275 GAATTTTTTTCTGGAGTTTTCTGTTCTCCAGTTTAAATGGTTCAGATGAGAGATGTAT 1334
Db 1334 GAATTTTTCTCTGGGTTTCTGTTCTCCAGCTTAAATTTGGTTCAGATGCGTATGTAT 1393
Qy 1335 TGAGCAGCTACAGCCAAATCAGAACTACTTTCCGGGCTGATGATGACACCAATTTGCTTA 1394
Db 1394 TGGGGCAGCAACAGCAATCAGAACTACTTTCCAGGCTGATGATGACCAATATCATTTGCTTA 1453
Qy 1395 CATTGAAATTTACTCCATTTCTTAACTTTGTCGAAAGCGAGTTCCGACTTGGTATGAATA 1454
Db 1454 CATTGAAATAATCTCTATTTCTCAGAGTGTGCAATTCGAGTTTCGGCTTGGGTTGAATA 1513
Qy 1455 TACATGGGACTCTCAAGAAATGCTAGATGATGCTGATTTGCTTTAAAGACCTTACCAACTAC 1514
Db 1514 TACATGGAACTCACAAGAAATCTAGATGATGCTCAACTTCTTCAGAACTTCCGGACAGC 1573
Qy 1515 GTCCAGTTAGCCCTCGCCATGATGTGAATCTTCAAGATCATCAGCAAAAGTTCGACTTGT 1574
Db 1574 AATGCACTTGTCTATTGCTTGTACATAAACTTCAGTATCATCGACAAGGTGGAGTTATT 1633
Qy 1575 CAAGGTTGTGATACACAGATGATTTATGATGATGCTTGAATTTGAATTCGGTCTCTTA 1634
Db 1634 CAAGGGCTGTGACACAGATGATTTATGACCTGCTCTGAATTTGAAATCCCACTATTTA 1693
Qy 1635 TTTGCTGTGTGACTTTGCTGCAAAAGGGAGAAATTTGGCAAGGAATGTATATCATCAA 1694
Db 1694 TTTACCTGGTGAATTTGCTGCAAAAGGGAGAAATTTGGAAGAAATGTATCATCAA 1753
Qy 1695 GCATGGAGAAGTCCAAGTCTTTGGAGCCCTGATGGTACTAAAGTTCTGTTACTCTGAA 1754
Db 1754 ACACGGAGAAGTCCAAGTCTTTGGAGCCCTGATGGTGTCTCAAGTTCTGTTACTCTGAA 1813
Qy 1755 AGCTGGGTGGTGTGAGAAATCAGCTTCTAGCAGCAGGAGGAGGAACCGTCCGAC 1814
Db 1814 AGCTGGGTGAGTGTGAGAAATTAAGCTTCTGCGCAAAAGGAGGAGGAATCGCCGGAC 1873
Qy 1815 TGCCAAATGTGTGGGCCCAAGGTTTGCATCTTTTAACTCTAGACAAAAGACCTTCCA 1874
Db 1874 AGCTGATGTTGTGGGCCCAAGGTTTGCATCTTTTAACTCTGACAAAAGACTTCCA 1933
Qy 1875 AGAAATTTCTAGTCATTTATCCAGATTTCTGAAGGATCTCTATGAAGAAAGCCAGAGTCT 1934
Db 1934 AGAAATTTCTGCTGATTTATCCAACTTCTTAAAGAGTCTCTCATGAAGAAAGCCAAATCT 1993
Qy 1935 TTTAAGCAGAGAGCTAAGCCGACAGAGCAACCCCTCCCAAGAAAGATCTTGGCCCTCT 1994
```

```
Db 1994 TTTAAGCCAGAGGGAAGACCACTCAGGCAATCCCTCGCAAGACCAGGACCTGCTTTCT 2053
Qy 1995 CTTCCACCGAAAGAGAGACACCCAACTGTTTAAACTCTCTCCTAGGAGGACACAGGAA 2054
Db 2054 TTTCCACCGAAAGAGAGACACCCCGAATGCTTAAAGTTCTCCTAGGAAACACAGGAA 2113
Qy 2055 AGCAAGTCTTGCAGAGACTACTCAAAATTGAAGCGAGAGCAAGAGCTTCAGAAGAAAGAAA 2114
Db 2114 GGTGGACCTTGAAGACTCTTTAAGGAAGAGAGAAACCAACCACTCAGAA----- 2164
Qy 2115 TTCTGAAGGAGGAGAGGAAGAAAGAAATGAAGATAAAACAAAAAGAAATGAAGA 2174
Db 2165 -----ATAAGGTGAAGTTCCAAAGGAAGAAG 2195
Qy 2175 TAAACAAAAGAAAATGAAGATAAAGGAAAAGAAAATGAAGATAAAGGAAAGAGA 2234
Db 2196 AAGAGAAAAGAAATATGAAGACAAAGGAACAGAGCTTAGCAGAAAAGAA----- 2246
Qy 2235 GCAGAGAGAGAGCCACTCGACAGACCTCGAATGTACAGCAAGTCTTATTTGCAGTGGAGGA 2294
Db 2247 -----AAAAAACCCCACTGGATAGTTCAAAGGCACACTAAGTTCTATTACAGTGGAGGA 2301
Qy 2295 AGAACCCCACTCAGTTTAGAAGGACAGTTTTTACCAGAGGAGCTTCTCGTCAATCACTCAT 2354
Db 2302 AATGTCCCAATCAATTAATAGGCACTTTTCCCAAGGACACATGCTCTATCAATTTGATCAT 2361
Qy 2355 TATCAGATGGTCTCTTCTGCTGAGGGCGGAGAGAGGTTCTTACTATTGAGTCAAGA 2414
Db 2362 CAGCAACATGG-TCCCTGCTACTGAGACTGGAGAAGAGATTCTGACTATCAAAAGTCAAGA 2420
Qy 2415 AAAGGCTAAGCAATAA 2430
Db 2421 AAGGCTAAGCAATAA 2436
```

RESULT 5  
LOCUS CQ716885 1350 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 2819 from Patent WO02068579.  
ACCESSION CQ716885  
VERSION CQ716885.1 GI:42277742

KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE 1  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of human exons or transcripts, for detecting expression and other uses thereof

JOURNAL Patent: WO 02068579-A 2819 06-SEP-2002;  
PE Corporation (NY) (US)

FEATURES  
source Location/Qualifiers  
1. 1350  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 35.1%; Score 853.4; DB 6; Length 1350;  
Best Local Similarity 99.9%; Pred. No. 2.6e-140;  
Matches 854; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGTTTAAATCGCTGACAAAGTCAACAAGTCAAGGCTATAGGAGAGAAACAATGAGAAT 60  
Db 1 ATGTTTAAATCGCTGACAAAGTCAACAAGTCAAGGCTATAGGAGAGAAACAATGAGAAT 60  
Qy 61 GAACAAAGTTCTCGTCCGAATGAAGAGGCTCTCACCCAAAGTAAATCAAGTCTCAGCAACC 120  
Db 61 GAACAAAGTTCTCGTCCGAATGAAGAGGCTCTCACCCAAAGTAAATCAAGTCTCAGCAACC 120

QY	121	ACAGCACGAGAGAAACAAAGGTGAAGAGAAATCTCTCAAAACCAAGTCAACTCCAGTC	180
Db	121	ACAGCACGAGAGAAACAAAGGTGAAGAGAAATCTCTCAAAACCAAGTCAACTCCAGTC	180
QY	181	AGTCTTGAAGAGCCACACACCAACATACAGACAAATCTCTCAAGAAATTCCTCTGGA	240
Db	181	AGTCTTGAAGAGCCACACACCAACATACAGACAAATCTCTCAAGAAATTCCTCTGGA	240
QY	241	GATCTGACCAAAACCCCTGACCCCTCAAAATGCAGCAGAAACCAACTGGAACTGCCAGAG	300
Db	241	GATCTGACCAAAACCCCTGACCCCTCAAAATGCAGCAGAAACCAACTGGAACTGCCAGAG	300
QY	301	CAGAGGAAATGAGACCCCGGGAAGAGGTGCAGCCAGCCACAAACCAAAACCCCTCTGCA	360
Db	301	CAGAGGAAATGAGACCCCGGGAAGAGGTGCAGCCAGCCACAAACCAAAACCCCTCTGCA	360
QY	361	GCTCTCTGTTATATAATGAGTATGCCGATGCCAGCTACACAACTGGTGAAGAGTACGCT	420
Db	361	GCTCTCTGTTATATAATGAGTATGCCGATGCCAGCTACACAACTGGTGAAGAGTACGCT	420
QY	421	CAAGAACAGCCCTCTCAAGAAAGTTGGTAGAGGAGATCTCTCTCACCCGAGGCC	480
Db	421	CAAGAACAGCCCTCTCAAGAAAGTTGGTAGAGGAGATCTCTCTCACCCGAGGCC	480
QY	481	AGCCCAAACTGCAAGGCCAGCGCTGTACACACAGTAAAGAAAGCGATGATAGCCA	540
Db	481	AGCCCAAACTGCAAGGCCAGCGCTGTACACACAGTAAAGAAAGCGATGATAGCCA	540
QY	541	ACAGAACTTACTACAGGCTGTGTGGTTCAAGTCAAAAGATGCTTTACACAGATAC	600
Db	541	ACAGAACTTACTACAGGCTGTGTGGTTCAAGTCAAAAGATGCTTTACACAGATAC	600
QY	601	TTAAAGCGAATTAACCTTCAAAACAGCATAGATTTCATACACAGATCGACTCTATCTCTG	660
Db	601	TTAAAGCGAATTAACCTTCAAAACAGCATAGATTTCATACACAGATCGACTCTATCTCTG	660
QY	661	TGGCTCTGCTGTGCTCTTGGCTATTAACCTGGAATGCTGGTTTATACACTGGCCCTC	720
Db	661	TGGCTCTGCTGTGCTCTTGGCTATTAACCTGGAATGCTGGTTTATACACTGGCCCTC	720
QY	721	GTCCTTCCCATATCAAAACCGCAGACATACACTACTGGCTTATTCGGACATCATATGT	780
Db	721	GTCCTTCCCATATCAAAACCGCAGACATACACTACTGGCTTATTCGGACATCATATGT	780
QY	781	GATATCATCTACCTTTATGATATGCTATTATCCAGCCAGACTCCAGTTTGTGAAGAGA	840
Db	781	GATATCATCTACCTTTATGATATGCTATTATCCAGCCAGACTCCAGTTTGTGAAGAGA	840
QY	841	GGAGACATAAATAGTG	855
Db	841	GGAGACATAAATAGTG	855

RESULT 6	
HUMCNCCB	
LOCUS	3025 bp DNA linear PRI 01-MAY-1995
DEFINITION	Homo sapiens clone hRNC2a retinal rod cyclic nucleotide-gated cation channel gene, complete cds.
ACCESSION	L15297
VERSION	L15297.1 GI:291914
KEYWORDS	cyclic nucleotide-gated cation channel; retinal protein.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Chen,T.Y., Peng,Y.W., Dhallan,R.S., Ahamed,B., Reed,R.R. and Yau,K.W.
TITLE	A new subunit of the cyclic nucleotide-gated cation channel in retinal rods
JOURNAL	Nature 362 (6422), 764-767 (1993)
PUBMED	7682292

REFERENCE	2 (bases 1 to 3025)
AUTHORS	Ahamed,B.
TITLE	Direct Submission
JOURNAL	Submitted (17-MAY-1993) Baheer Ahamed, Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, MD 21205, USA
COMMENT	Original source text: Homo sapiens DNA.
FEATURES	Location/Qualifiers
source	1..3025
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/clone="hRNC2a"
	/tissue_type="retinal"
	589..2460
CDS	/codon_start=1
	/product="cyclic nucleotide-gated cation channel"
	/protein_id="AAA65619.1"
	/db_xref="GI:790520"
	/translation="MLCKFKRPMKKYQFQSDIPDLNLMYVLMFFVVMANNWN("LIPVWAPFYQTPDNIIHMLMDYLDLIYFLDITVQTRIQFVRRGGIITOKKIMRN NLYKRRFKMDLLSLPLDLYLVKGVNPLRLPCLKYMAFFEFNSRLSILSKAYV YRVIRTTALLVSLHNSCLYWSAYOGLGSTHMVYDGVNSYIRCYFYAVKTLITL GGLDPKTLPEIVFQLLNYFTGVFAFSVMIGMRDVGAATAGOTYYRSCMDSTVKYM NFKIPKSVQNRVKTWYETVHWSQMLDESELMVOLPKMRDLDAIDVYNIIVSKVAL FQGDQRMIFDMLKRLRSVYLPNDYVCKGEIGREMYIIQAGOVQVILGGPDGKSLV TLKAGSVFGEISLLAVGGNRRNTANVAHGTNLFILDKDLNELVHYPSQKLRK KARMLRNNKPKBEKSVLILPPRAGTPKLFNAALAMTGMKGKAKGKLAHLRRL KELAALAAAKHEELVEOKSSQDVKGEESAAPODHTHPKEAATDPAPRTPPPPG SPFSPFPASLGSCGESEGPABPEHSVRICMSPGPEPGEQILSVKMPEREKAE"
ORIGIN	

Query Match	23.8%;	Score	577.4;	DB	8;	Length	3025;
Best Local Similarity	64.3%;	Pred. No.	8.1e-92;				
Matches	866;	Conservative	0;	Mismatches	481;	Indels	0;
Gaps	0;						
QY	617	TTCCAAAACAGCATAGATT	CATACACAGATCGACTCTATCTCTGTGGCTCTTGGCTGTCA	676			
Db	635	TTCCCCAGAGCATTCACCCCGCTGACCAACCTGATGTATGCTCTATGGCTGTCTTCGTGG	694				
QY	677	CTCTTGCCTATAACTGGAACTGCTGGTTTATACCACTCGCCCTCGCTTCCCATATCAAA	716				
Db	695	TGATGGCTTGGAACTGTGGAACTGTGTGGCTGATTTCCCGTGGCTGGCCCTTCCCTTACCAGA	754				
QY	737	CCGCAGACAACATACACTACTGGCTTATTGGCGACATCATATGTATCATCTACCTCTT	796				
Db	755	CCCCGACAACATCCACCACCTGGCTGCTGATGGATTACCTATGCGACCTCATCTACTTCC	814				
QY	797	ATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTGAAGAGGAGAGACATAATAGTGG	856				
Db	815	TGGACATCACCGTGTTCAGACACACGCTCGAGTTTGTCTAGAGGCGGGGACATCATTCACG	874				
QY	857	ATTCAAATGAGCTAAGGAAACACTACAGACTTCTACAAATTTTCAGTTGGATTTGGAT	916				
Db	875	ACAAAAGGACATGCGAAATACCTACCTGGAAGTCTCGCCGCTTCAAGATGGAACTTATTA	944				
QY	917	CAATAATACCAATTTGATATTTGCTACCTCTTCTTTGGGTTTAAATCCAAATGTTTAGAGCAA	976				
Db	935	GCCTCTGCCCTTGGATTTTCTCTATTTCAAAAGTCGGTGTGAACCCCTCTCCCGCTGC	994				
QY	977	ATAGGATGTTAAGTACACTTCATTTTTTGAATTTAATCATCATCCCTAGAGTCTATAATGG	1036				
Db	995	CCCGCTGTTTAAAGTACATGCGCTTCTTCGAGTTTAAACAGCGCTTGGAAATCCATCTCA	1054				
QY	1037	ACAAAGCATATATCTACAGAGTATTTCGAACAACCTGGATCTGCTGTTTATTTCTGCACA	1096				
Db	1055	GCAAGCCTACGTGTACAGGGTCATCAGGACCAAGCCCTACCTTCTCTACAGCCTGCATT	1114				
QY	1097	TTAATGCCCTGTGTTTATTACTGGGCTTCAAACTATGAAGGAATTTGGCACTACTAGATGG	1156				
Db	1115	TGAATTCCTGCTTTATTACTGGGCATCGGCTATCAGGGCTCTCGCTCCACTCACTGGG	1174				
QY	1157	TGTATGATGGGAAGGAAACGAGTATCTCAGATGTTATTATTTGGGCAGTTCGAATTTAA	1216				





misc_feature		/gene="CNCG2" /note="672 bp PCR product" /citation=[2] 1701..4033 /gene="CNCG2" /note="cDNA clone hRNC2a-like portion" /citation=[2] /function="modulates alpha subunit of the cGMP-gated cation channel" 3814..4033 /gene="CNCG2"	
3' UTR			
ORIGIN			
Query Match		23.8%; Score 577.4; DB 8; Length 4033;	
Best Local Similarity		64.3%; Pred. No. 7.5e-92;	
Matches		866; Conservative 0; Mismatches 481; Indels 0; Gaps 0;	
QY	617	TTCCAAACAGCATAGATTTCATACACAGATCGACTCTATCTCCCTGCTGCTCTTGTGCA	676
DB	1988	TTCCCCAGAGCATGACCCGCTGACCAACCTGATGATGCTTATGGCTGTTCTTCGTGG	2047
QY	677	CTCTTGCTATACTGGAATCTGTTTATACCACTGCGCCTCGTCTTCCCATATCAAA	736
DB	2048	TGATGGCTGGAATTTGGAATCTGCTGATTCCTGCTGCTGGGCTTCCCTACAGA	2107
QY	737	CCGACAGACATACACTACTGCTTATTCGGACATCATATGTATATCATCTACTTT	796
DB	2108	CCCGGACACATCCACCACTGCTGCTGATGATTAACCTATGCGACCTCATCTACTTC	2167
QY	797	ATGATATGCTATTATCCAGCCAGACTCCAGTTTGTAGAGGAGGAGACATATAGTGG	856
DB	2168	TGGACATACCCGTTTCCAGACACGCTGCGATTGTGAGGGGGGACATCATACGG	2227
QY	857	ATTCAATGAGCTAAAGAAACACTACAGGACTTCTACAAAATTTTCAGTTGGATGTCGCAT	916
DB	2228	ACAAAAGGACATGCGAAATAACTACCTGAAGTCTGCCGCTTCAAGATGACCTGTCTCA	2287
QY	917	CAATAATACCATTTGATTTGTTGCTTCTTCTTTGGGTTTAATCCAAATGTTTAGAGCAA	976
DB	2288	GCCTCCTGCTTGGATTTTCTCTATTGAAAGTCGGTGTGAACCCCTCCTCGCGCTGC	2347
QY	977	ATAGGATGTTAAGTACACTTCTTTTTCGAATTAATCATCATCAGCTAGAGTCTATATGG	1036
DB	2348	CCCGCTGTTTAAAGTACATGGCTTCTTCGAGTTTAAACGCCGCTCGAATCCATCTCA	2407
QY	1037	ACAAAGCATATATCTACAGAGTTATTTCGAAACAACTGGATATCTGCTGTTTATTCTGCACA	1096
DB	2408	GCAAAGCTACGTGTACAGGGTCACTACGAGCACAGCTACTTCTCTACAGCTGCAATT	2467
QY	1097	TTAATGCTGTTTATTACTGGGCTTCAAACTATGAAGAAATGGCACTACTAGATGGG	1156
DB	2468	TGAATTCCTGTTTATTACTGGGCATCGGCTATCAGGGCTCGGCTCCACTCACTGGG	2527
QY	1157	TGATGATGGGAAGGAACGAGTATCTGAGATGTTTATTATGGGCACTTGCACCTTAA	1216
DB	2528	TTTACGATGGCGTGGAAACAGTTTATTCGCTGTACTACTTTGCTGTGAAGACCTCA	2587
QY	1217	TTACATTTGGTGGCTTCCAGAACCAAACTTTATTGAAATGTTTTCAACTCTTGA	1276
DB	2588	TCACATTCGGGGGCTGCTGACCCCAAGACACTCTTTGAAATGTTCTTCAGCTGCTGA	2647
QY	1277	ATTTTTTTCGAGTTTTTGTGTTCTCCAGTTTAAATGTTGTCAGATGAGAGATGATG	1336
DB	2648	ATTATTTTCAGGGGCTTCTTCTTCTGATGATCGGACAGATGAGAGATGCTGTAG	2707
QY	1337	GAGCAGCTACAGCCAAATCAGAACTACTTCCGGCTGCTGATGATGACACCAATTGCCTACA	1396
DB	2708	GGGCGCCACCGCGGACAGACCTACTACCCAGCTGCAATGGACAGCACGGTGAAGTACA	2767
QY	1397	TGAACAAATTAATCTCAATCTTAACTGTGCAAAAGCGAGTTTCGCACTTGGTATGAATATA	1456
DB	2768	TGAATTTCTACAGATCCCAAGTCCGTCGAGAACCCGCTCAAGACCTGGTACGATACA	2827
QY	1457	CATGGACTCTCAAGAATGCTAGATGAGTCTGATTTTGCTTAAAGCCCTACCAACTACGG	1516
DB	2828	CCTGGACTCGCAAGGCATGCTGGATGAGTCAAGCTGATGCTGCAGCTTCCAGACACAGA	2887
QY	1517	TCAGTTAGCCCTCGCCATTGATGTGAACCTTCAGCATCATCAGCAAGTCGACTTGTTC	1576
DB	2888	TGCGGCTGACCTCGCCATCGACGTGAACCTACAACATCGTTAGCAAGTCGCACCTTTTC	2947
QY	1577	AGGTTTGATACACAGATGATTTATGACATGTTGCTAAGATTGAAATCCGTTCTCTATT	1636
DB	2948	AGGCTGTGACCGGCGAGATGATCTTTGACATCTGAAGAGGCTTCGCTCTGTGCTACC	3007
QY	1637	TCCTGGTGAATTTGTCTCAAAAAGGGAGAAATTTGGCAAGGAAATGTATATCATCAAGC	1696
DB	3008	TGCCCACACATATGTTGTCAGAGAGGGGAGATCGGCCGTGAGATGTACATCATCCAGG	3067
QY	1697	ATGGAGAACTCCAAGTTCTTGGAGCCCTGATGTGTAATAAGTTCTGGTTACTCTGAAAG	1756
DB	3068	CAGGCAAGTGCAGGTCTTTGGGCGGCCCTGATGGGAAATCTGTGCTGGTGACGCTGAAAG	3127
QY	1757	CTGGTCTGTTTGGAGAAATCAGCCTTCTAGCAGCAGGAGGAGAAACCGTTCGAACTG	1816
DB	3128	CTGATCTGTGTTTGGAGAAATAAGCTTGTGCTGCTGGGAAATCTGTGCTGGTGACGCTGAAAG	3187
QY	1817	CCAACTGCTGCCACCGGTTTGGCAATCTTTTAATCTCTAGACAAAAGACCTCTCAAG	1876
DB	3188	CCAACGTGTGGCGCACGGGTTTACCACTCTTCTCAGAGTTACTTCCGGAAGAAAGCCAGCGATGC	3247
QY	1877	AAATTTAGTGCATTTATCCAGATTTCTGAAAGGATCTCTATGAAGAAGCCAGAGTGTCTT	1936
DB	3248	AGATTTTGTGCAATTATCTGAGTCTCAGAGTTACTTCCGGAAGAAAGCCAGCGATGC	3307
QY	1937	TAAAGCAGAGGCTTAAGCCGCAAG 1963	
DB	3308	TGAGAAGCAACAATAAGCCCAAGGAG 3334	
RESULT 9			
AF042498			
LOCUS			
DEFINITION			
Homo sapiens rod photoreceptor CNG-channel beta subunit (RNCN2)			
mRNA, complete cds.			
ACCESSION			
AF042498			
VERSION			
AF042498.1 GI:2921582			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
IDENTIFICATION of a domain on the beta subunit of the rod			
cGMP-gated cation channel that mediates inhibition by			
calcium-calmodulin			
J. Biol. Chem. (1998) In press			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1..4382			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
1..4382			
/gene="RNCN2"			
71..3808			
/gene="RNCN2"			
/note="cyclic nucleotide-gated cation channel beta			
subunit"			
/codon_start=1			

/product="rod photoreceptor CNG-channel beta subunit"  
/protein\_id="AAC04830.1"  
/db\_xref="GI:2921593"  
/translation="MLGMVQVLPPOPGRPRKTKMQEVEEVEPEMEAEVPEPNE  
EAEVSEWPPPESEKBEVAVADPSQPTKEAALTSTLSRAQAEISEMNSPSHRV  
LTLMLKGVKVPQPVHSTEDPAQILGHSGTDGCTDEPNEAEQDTRFGLRL  
WLEQLERVLPPPKSSSEWRDEPAVATAPPGQPMGPKQARETSPSLPIQPK  
BEPKAPAPQPGSQATSLPPTRDPAVLVHLRLEMAALPOVHLKIGESEP  
SPGICDQVTSILPGQVPEPDLVBEVEPPWEDAHQDVSTSPQTEVPVPAVEENKAV  
EKMPRELRSIEBEKEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEE  
STSDKLEVEEKEAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE  
GVPAKQHPQVQVQDADSCPLMAEENPSTVLPPSPAKSDTLIVPSSASGTHKK  
LPSEDEABELKALSPASPVVAVSDPTPKTDGQDRAASTASTASINLRQLELV  
KLFKERTKVEKLLIDPDVTSDESPKSPAKKAPAPADPTKPAEAEVPEEHEYCDML  
CKKFRHPWKYQFQSDIDPLNLMVLMFFVVMAMNWCWLIIVRWAFPYQTDNI  
HMLMLDYLCDLIYFLDITVFTQLQFVRGGDIITDKDMRNLYLKSRRFKMDLSLL  
PLDLYLVKGVNPLRLPCLKVAFFENSRLESLSKAVYVRVIRVITAYLLYSHL  
NSCLYMASAYOGLGSTHWVDGNSVTRCYFAVKTLITIGGLDPEKTLPEIIVFQL  
LNYFTGVFAPSVNIGMRDVGAATAGQTYRSCMDSTVKVYNNFKIPKSVQNRKWT  
YBYTHSQMLDESELMQDRLDLDLADVNYNIVSKVALFQGCDDQMIFDMLKRL  
RSVYLPNDYVCKGEIGREMYIIQAGQVQLGPDGKSLVTLKAGSVFGEISLLAV  
GGGNRTANVAGHTNLFILDKDLINELVHPESQKLLRKKARMLRSNNKPKEEK  
SVLLPFPRAQTPKLNAALMTGKMGKGKGLAHLRLRLKELAALEAAKQOELV  
EQAKSQDVKGEGSNAAPDQTHPKAATDPPAPTPPEPPSPSSPPASLGRPEG  
EBEGAPAEHESVRICMSPGPEPQILSVKMPEREKAE"

ORIGIN

Query Match		23.8%; Score 577.4; DB 8; Length 4382;
Best Local Similarity		64.3%; Pred. No. 7.3e-92;
Matches 866; Conservative 0; Mismatches 481; Indels 0; Gaps 0;		
QY	617	TTCCAAAACAGCATAGATTCAATACAGATCGACTCTATCTCTGTGGCTCTTGGCTGTGCA 676
DB	1983	TTCCCGCAGAGCATTGACCCGCTGACCACTGATGTATGCTATGCTGTGGCTGTCTTCGTGG 2042
QY	677	CTCTTGGCTTAACTGGAAGTCTGGTTATACCAGTGGCTCTGCTTCCCATATCAAA 736
DB	2043	TGATGGCTTGGAAATGGAACTGTGGCTGATTTCCCGTGGCTTCCCTTACCAGA 2102
QY	737	CGCAGACAACATACACTAGCTGGCTTATTCGGAGCATCATATGTGATATCATCTACCTTT 796
DB	2103	CCCCGGACAACATCCACACTGGCTGCTGATGATTAACCTATCGACCTCATCTACTTCC 2162
QY	797	ATGATATGCTATTTATCGACCCAGACTCCAGTTTGTAAAGAGAGAGACATAAATAGTGG 856
DB	2163	TGGACATCACCGTGTTCAGACACGCGCTGCAGTTTGTTCAGAGCGCGGGACATCATTCGG 2222
QY	857	ATTCAAATGAGCTAAGGAAACACTACGAGCTTCTACAAATTTTCAGTTGGATGCGCAT 916
DB	2223	ACAAAAGGACATGCGAAATACTACTGAAATGTCGCGCGCTTCAAGTTCAGACCTGCTCA 2282
QY	917	CAATAATACCAATTTGATATTTGCTACTCTTCTTTGGGTTTAAATCCAATGTTTGAAGCAA 976
DB	2283	GCCTCTCGCTTGGATTTTCTATTTTGAAGTCGGTGTGAACCCCTCTCCGCGCTGC 2342
QY	977	ATAGATGTTAAAGTACACTTCATTTTGAATTTAATCATCACCTAGAGTCTATAATGG 1036
DB	2343	CCCGCTGTTTAAAGTACATGGCTTCTTCCAGTTTAAACAGCGCGCTTGAATCCATCCTCA 2402
QY	1037	ACAAAAGCATATATCTACAGAGTTATTCGAACAACATGGATCTGCTGTTTATCTCGACA 1096
DB	2403	GCAAAAGCTTACGTACAGGGTCATCAGACCACAGCCTTACCTTCTTACAGCCCTGCAAT 2462
QY	1097	TTAATGCCCTGTGTTTATTACTGGGCTTCAAACTATGAAGAAATGGCACTACTAGATGGG 1156
DB	2463	TGAAITCTCTGTTTATTACTGGCATCGGCTATCAGSGCCTCGGCTCCACTCTACTGGG 2522
QY	1157	TGTATGATGGGGAAGAAACGAGTATCTGAGATGTTTATTATTTGGGAGTTCGAACTTTAA 1216
DB	2523	TTTACGATGGCGTGGGAAACAGTTTATTTTCGCTGTACTACTTCTTACAGCCCTGCAAT 2582
QY	1217	TTACCAITGGTGGCTTCCAGAACACAACTTTATTGAAATGTTTTTCACTCTTGA 1276

DB	2583	TCACCATCGGGGGCTGCTGACCCCAAGACACTCTTTGAAATTTCTTCTCCAGCTGCTGA 2642
QY	1277	ATTTTTTTTCTGGAGTTTTTGTGTTCTCCAGTTTAATTTGGTCAGATGAGAGATTTGATTG 1336
DB	2643	ATTATTTTCCGGGCGCTCTTTGCTTCTCTGTGATCATCGGACAGATGAGAGATTTGGTAG 2702
QY	1337	GAGCAGCTACAGCCAAATCAGAACTACTTTCGCGCCTGTGATGATGACACCACTTGGCTACA 1396
DB	2703	GGCGCCCAACCGCGGACAGACCTACTACTCCGAGCTGATGACACGACCGGTGAAGTACA 2762
QY	1397	TGAACAATTAATCCATTCTCTAACTTTGTGCAAAAGCGAGTTTCGAGTTTGGTATGAATATA 1456
DB	2763	TGAATTTCTACAAGATCCCAAGTCCGTCGACAAACCGGTCAAGACTTGGTACGAGTACA 2822
QY	1457	CATGGGACTCTCAAAAGAAATGCTAGATGATCTGATTTTCTTAAAGACTTTTAACTTTTAA 2882
DB	2823	CCTGGCACTCGCAAGGCATGCTGGATGATGAGCTGATGCTGAGCTTCCAGAACACA 2942
QY	1517	TCCAGTTAGCCCTCGCCATTTGATGTGAATCTCAGCATCATCAGCAAAAGTCGATTTGTTCA 1576
DB	2883	TGCGCTGGACCTCGCCATCGACGTGAATCAACAATCGTTAGCAAAAGTCGCACTCTTTC 2942
QY	1577	AGGTTGTGATACACAGATGATTTTATGACATGTTGCTAAGATTGAAATCCGTTCTCTATT 1636
DB	2943	AGGCTGTGACCGGCAGATGATCTTTTGACATGCTGAAGAGGCTTCGCTCTGTGTTCTACC 3002
QY	1637	TGCTGTGCTACTTTGCTGCAAAAGGAGAAATTTGGCAAGGAAATGTATATCATCAAGC 1696
DB	3003	TGCCCCAAGCATATGTTGCAAGAGGGGAGATCGGCGGTGAGATGTACATCATCCAGG 3062
QY	1697	ATGAGAAAGTCAAGTCTTGGAGCCCTGTATGTACTTAAAGTTCTGTTACTCTGAAAG 1756
DB	3063	CAGGCAAGTGCAGTCTTGGCGGCCCTGTATGGAAATCTGTGCTGTGACGCTGAAAG 3122
QY	1757	CTGGTCTGCTTTTGGAGAAATCAGCTTCTTAGCAGCAGGAGGAGAAACCGTCCAACTG 1816
DB	3123	CTGGATCTGTGTTTGGAGAAATAAGCTTGTGGGTGTTGGGGGGGGAACCGGCGCACGG 1882
QY	1817	CCAATGTGCTGGCCCAACGGGTTTGGCAATCTTTTAACTCTAGACAAAAAGACCTTCCAAG 1876
DB	3183	CCAAGTGTGGCGCACGGTTTACCAACCTTTCATCTCGATAAAGAACCTGTAATG 3242
QY	1877	AAATTTCTAGTGATATCCAGATTTGAAAGATCTCATGAAGATCCCTCATGAAGACGAGTGT 1936
DB	3243	AGATTTGTGATATCTCTGAGTCTCAGAAATTTACTCCGAAAGAAAGCCAGGCGCATGC 3302
QY	1937	TAAACGACAGGCTAAGACCCGACAGAG 1963
DB	3303	TGAGAACACAATAAGCCCCAAGGAGG 3329

RESULT 10

LOCUS	CQ715550	4320 bp	DNA	linear	PAT 01-FEB /004
DEFINITION	Sequence 1484 from Patent WO02068579.				
ACCESSION	CQ715550				
VERSION	CQ715550.1				
KEYWORDS	GI:42276407				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.				
TITLE	Kits, such as nucleic acid arrays, comprising a majority of human exons or transcripts, for detecting expression and other uses thereof				
JOURNAL	Patent: WO 02068579-A 1484 06-SEP-2002;				
FEATURES	PE Corporation (NY) (US)				
source	Location/Qualifiers				
	1..4320				
	/organism="Homo sapiens"				

Query Match	23.7%	Score 575.8;	DB 6;	Length 4320;
Best Local Similarity	64.2%;	Pred. No. 1.4e-91;		
Matches 865;	Conservative 0;	Mismatches 482;	Indels 0;	Gaps 0;
Qy	617	TTCAAACAGCATAGATTTCATACACAGATCGACTCTATCTCTGTGGCTCTTGCTTGTCTCA	676	
Db	1928	TTCCCAGAGCAATGACCCGCTGACCAACTGATGTATGTCTATGCTTCTTCTCGTGS	1987	
Qy	677	CTCTTGCTATAACTGGAACCTGCTGGTTATACACTTGGCCCTGCTTCTCCCATATCAAA	736	
Db	1988	TGATGGCTTGGAAATGGAATGTTTGGCTGATTCCTGGCTGGGCTTCCCTACACAGA	2047	
Qy	737	CCGACAGCAATACACTACTGCTTATTCGGGACATCATATGTGTATATCATCTACTCTTT	796	
Db	2048	CCCGGACAAATCCCACTGCTGCTGATGTATCTTACCTATGCGACCTCATCTACTCTCC	2107	
Qy	797	ATGATATGCTATTTATCCAGCCAGACTCCAGTTTGTAAAGAGGAGACATAAATAGTGG	856	
Db	2108	TGGACATACCGTGTTCAGACACAGCCCTGCAGTTTGTACAGGCGGGACATCATTAACGG	2167	
Qy	857	ATTCAAATGAGCTAAGGAAACATACAGGACTTCTACAAATTTTCAGTTGGATGTCGCAT	916	
Db	2168	ACAAAAAGGACATCGGAAATAACTACCTGAAGTCTCGCCGCTTCAAGATGGACCTGCTCA	2227	
Qy	917	CAATAATACCAATTTGATATTTGCTTACCTCTCTTTTGGTTTAAATCCAATGTTTAGAGCAA	976	
Db	2228	GCCTCTGCCCTTGGATTTTCTCTATTTGAAAGTCGGTGTGAACCCCTCTCCGCGCTGC	2287	
Qy	977	ATAGGATGTTAAAGTACACTTCAATTTTGTAAATTAATCACTAGAGTCTATAATGG	1036	
Db	2288	CCGCTGTTTAAAGTACATGCGCTTCTTCAGTTTAAACAGCGCTGGAAATCCATCTCTCA	2347	
Qy	1037	ACAAAGCATATCTACAGAGTTATTCGAACAACCTGGATCTTCTGTGTTTATTTCTGCACA	1096	
Db	2348	GCAAGGCTACGTGTACAGGCTCATCAGGACCAAGCTTCTCTTCTCAGCGCTGCATTT	2407	
Qy	1097	TTAATGCTGTGTTTATTTACTTGGGCTTCAAACATGAAGNAATGGCACTACTAGATGGG	1156	
Db	2408	TGAATTCCTGCTTTTACTTGGGCTCGGCTTATCAGGCGCTCGCTCACTCACTGGG	2467	
Qy	1157	TGATGATGGGGAAGGAAAACAGATATCTGAGATGTATTAATTTGGGCAAGTTGGAACITTA	1216	
Db	2468	TTTACGATGGCGTGGGAAACAGTTATATTCGCTGTACTACTTTTGTGTGAAGACCTCA	2527	
Qy	1217	TTACCATTTGGTGGCTTCCAGAACCAAACTTTATTTGAAATGTTTTTCAACTCTTGA	1276	
Db	2528	TCACATCGGGGGCTGCTGACCCCAAGACACTCTTTTGAATTTGCTTCCAGCTGCTGA	2587	
Qy	1277	ATTTTTTTCTGGAGTTTTTGTGTTCTCCAGTTTAAATTTGTTTCAGATCAGAGATGTGATG	1336	
Db	2588	ATTATTTACGGCGCTCTTTGCTTCTGTGTATGATCGGACAGATGAGAGATGTGCTAG	2647	
Qy	1337	GAGCAGCTACAGCCAAATCAGAACTACTTCCGGGCTGCATGGATGACACCATTCGCTTACA	1396	
Db	2648	GGGCGGCACCCGACGAGACAGACCTACTACCGCAGCTGCATGGACAGCACCGGTGAAGTACA	2707	
Qy	1397	TGAACAAATTAATCTCAATTCCTAAAATTTGTGCAAAACGAGTTTCGGACTTGGTATGAATATA	1456	
Db	2708	TGAATTTCTACAAGATCCCAGGTCGGTGCAGAACCCGGTCAAGACCTGGTACGAGTACA	2767	
Qy	1457	CATGGGACTCTCAAAGAATGCTAGATGATCTGATTTGCTTTAAGACCCCTACCAACTACGG	1516	
Db	2768	CCTGGCACTCGCAGGCAATGCTGGATGATGATCAGAGCTGATGGTCGAGCTTCCAGACAAGA	2827	
Qy	1517	TCCAGTTAGCCCTCGCCATTCGATGTGAATTTTCAGCATCATCAGCAAAAGTTCGACTTTTCA	1576	
Db	2828	TGCGGCTGGACCTCGCCATCGACGTGAATCAACATCGTTAGCAAAAGTTCGACTCTTTC	2887	
Qy	1577	AGGGTGTGTATACACAGATGATTTATATGACATGTTGTGAAGATTTGAAATTCGGTCTCTATT	1636	

Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: nisc\_mgc@hghri.nih.gov  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granice, S., Guan, X., Gupta, J., Haghghi, P.,  
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaspi, R.,  
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Skatripop, S., Thomas, P.J., Touchman, J.W.,  
 Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAK plate: 85 Row: 1 Column: 9.

#### FEATURES

Location/Qualifiers  
 1..4763  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4504353"  
 /tissue\_type="Eye, retina, mouse strain C57Bl/6"  
 /clone\_lib="NIH MGC\_94"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

#### gene

<1..4763  
 /gene="Cngb1b"  
 /note="synonym: Cngb1"  
 /db\_xref="GeneID:33329"  
 /db\_xref="MGI:2664102"  
 <1..2602

#### CDS

/gene="Cngb1b"  
 /codon\_start=2  
 /product="Cngb1b protein"  
 /protein\_id="AAH45114.1"  
 /db\_xref="GI:28175675"  
 /db\_xref="GeneID:33329"  
 /db\_xref="MGI:2664102"  
 /translation="EETPIVLDSCLVQVADVDEQLBTPSELASIOELPEKEEKE  
 BEKEBEKEEKEEVEKEEGEATSTVPATKEHPELOVEDTDADGGPLIPEETL  
 PPPPPPPSVKSDTLTPVGAAGHKKLPSQDEASELKALESAPSVVAVMSDPTT  
 POEADQDPAASTAQNSAI INDRLOELVMEKTERKVEKLIIDPDVTSDEESKPS  
 FAKKAPEDPAQPAEAEVEBEHYCDMLCCFKRRPLKMYRFPPOSIDPLNLMYLW  
 LFFVVLWNWNCWLIIPVRWAFYPQADNIHFWLLMDYLCDFIYLLDIITVFQMLQFVK  
 GGDITTDKEMNNYTKRRFRMDLCLLPDLFLYLKGINPLRLPRCLKYMAFPEF  
 NNRLBAIILSKAYVYVIRTYLLVLSHLNSCLYVWAFQIGSTHWYDGVNSYI  
 RCYVWAKTLITIGGLPDPOTLFEIVFOLLNYFTGVFAFSVMIGOMRDVVGATAGOT  
 YVRSMDSTVKMYNPKYIPRSVQNRVKTWYETHWSQGMDESELMVQLPKDRULDLA  
 IDVNTSIVSKALPFGCDQRMIFDMLKRLRSVILPNDYVCKKGIGREMYIIQAGV  
 LVHPESDQKAVLITKAGSVFGEISLLAVGGNRRNTANVAHGFNLFILOKOLNEI  
 LVHPESDKLRLKARRLNNNKPEKESVLILPPRAGTPKLFNAAALAAAGKMPRG  
 AKGGKLHLRLKELALEAAARQOLLEQAKSQEAGGEGSGATDQAPAPBPPEP  
 KDPPKPPGPPPSAQSSPPPSAKFEESTGEAAGPEPSVRIRVSPGPDPEQTLSVE  
 VLEKEKEGAE"

#### ORIGIN

Query Match 23.7%; Score 575.2; DB 9; Length 4763;  
 Best Local Similarity 63.6%; Pred. No. 1.7e-91;  
 Matches 915; Conservative 0; Mismatches 508; Indels 15; Gaps 2;  
 QY 617 TTCCAAACGAGTAGATTATACACAGATCGACTCTATCTCTGTGGCTTCTGTGCA 676  
 DB 783 TCCCCAGAGTATGACCCACTGACCACTCATGTATACATCTTGTGGCTGTCTTTGTGG 842  
 QY 677 CTCTTGGCTTAACTGGAACTGCTGGTGTATATACCATCGGCTCGCTTCCCATATCAA 736  
 DB 843 TGCTGGCTTGGAACTGGAAGTATCCCTGTGGCTGTGGCTTCCCGCTTCCCGTACCAGC 902

QY 737 CCGCAGACAAACATACACTACTGGCTTATTGGCGACATCATATGTGATCATCTACCTTT 796  
 DB 903 GGGCAGACAAACATCCACTTCTGGCTCCTCATGGATTACTTGTGGCACTTCTACTACCTCC 962  
 QY 797 ATGATATGCTATTTATCCAGCCAGACTCCAGATTGTAAGAGGAGGAGACATAATAGTGG 856  
 DB 963 TGGACATCACCGTGTTCAGATGCGCTTGCAGTTTGTCAAAGCGGGGACATCATTAGAG 1022  
 QY 857 ATTCAAAATGAGTAAGGAAACACATACAGACTTCTACAAAATTTTCAGTTGGATGTCGCAT 916  
 DB 1023 ACAAGAGGAGATGCGTAATAACCTGAAAGTTCGCGCGTTTAAAGATGACCTGCTCT 1082  
 QY 917 CAATAATACCAATTTGATATTTGCTTACCTCTCTTTGGGTTTAAATCCAATGTTTACAGCAA 976  
 DB 1083 GCCTCTGCGCTTGGACTTTCTACTTTGAAACTTGGCATCAACCTCTCTCTCGCCTGT 1142  
 QY 977 ATAGATGTTTAAAGTACACTTCACTTTTGAATTTTAAATCATCATCATAGAGTCTATAATG 1016  
 DB 1143 CCCGCTGCTGAAGTACATGCGCTTCTTTGAGTTTAAATAACCGCTTGAAGCCATCTCTCA 1202  
 QY 1037 ACAAGCATATATCTACAGAGTTATTGCAACAACTGGATACTTGTCTGTTTATCTGCAACA 1096  
 DB 1203 GCAAAGCTTACGTTTACAGAGTTCATCAGGACCAAGCGCTTACCTGTGTACAGCTGCAAC 1262  
 QY 1097 TTAATGCTGTGTTTATTACTGGGCTTCAAACTATGAAGAAATTTGGCACTACTAGATGGG 1156  
 DB 1263 TCAACTCTGCTTTTACTTCTGGGCATCAGCTTCCAGGCGCATCGGCTTCCACTACTGGG 1322  
 QY 1157 TGTATGATGGGGAAGGAAACGAGTATCTGAGATGTTTATTATTTGGGCACTTTCGAACTTTAA 1216  
 DB 1323 TTTACGACGAGTGGGGAACAGTTTATATTCGATGCTACTTCTGCTGTGGCTGTGAAACCTCTA 1382  
 QY 1217 TTACCATTTGGTGGCTTCCAGAACCAACAACTTTATTGAAATTTCTTTTCAATCTTTCA 1276  
 DB 1383 TCACCATCGGAGGACTGCCCGACCCAGACGCTCTTTTGAGATTGTCTTCCAGCTGTCTGA 1442  
 QY 1277 ATTTTTTTCTGGAGTTTTTGTGTTCTCCAGTTTAAATTTGGTCAGATGAGAGATGTGATTG 1334  
 DB 1443 ACTATTTTACCGGGGTCTTTGCTTCTCTGTGATGATTGGACAGATGAGAGATGTGTTGG 1502  
 QY 1337 GAGCAGCTACAGCCAACTCAGAACTACTTTCGCGGCTTGCATGATGATGACCACTTGTCTACA 1394  
 DB 1503 GGGCGCCACACGCGGGGAGACCTACTACCGCAGCTGCATGACAGACAGTGAAGTACA 1562  
 QY 1397 TGAACAAATTTACTCCATTTCTTAACTTTGTGCAAAAGCGAGTTTCGGACTTGGTATGAATA 1456  
 DB 1563 TGAATTTCTCAAGATCCCAAGTCTGTGCAAGAACCGGTGTCAAGACCTGGTATGAGTACA 1622  
 QY 1457 CATGGGACTCTCAAGAAATGCTAGATGAGTCTGATTTGCTTTAAGACCCCTACCAACTACGG 1516  
 DB 1623 CTTGSCATTCGCAAGGCATGCTGGATGAGTCAGAGCTGATGTTCAACTTCCGACACAAGA 1682  
 QY 1517 TCCAGTTAGCCCTCGCCATCATGTGAACTTTCAGCATCATCAGCAAAAGTCGACTTGTTC 1576  
 DB 1683 TCGCACTGGACTCGCCATTGATGTGAATACAGTATTTGTGAGCAAAAGTGGCCCTCTTTTC 1742  
 QY 1577 AGGGTTGTGATACAGATGATTTATGACATGTTGCTTAAGATTGAAATCCGCTTCTCTATT 1636  
 DB 1743 AGGGCTGTGACCGGCAGATGATCTTCGACATGCTCAAGCGGCTTCTGTTCTGTCTTACC 1802  
 QY 1637 TGCCTGTGACTTTGCTGCAAAAAGGAGAAATTTGGCAAGAAATGTATATCATCAAGC 1696  
 DB 1803 TACCAATGACTATGTGTGCAAGAGGAGGAGATTTGGCCGAGAGATGTATCATATCCAGG 1862  
 QY 1697 ATGAGAAAGTCCAAAGTTCTTGGAGGCCCTGATGTTACTTAAAGTTCTTGGTTACTCTGAAAG 1756  
 DB 1863 CGGGCAGGTGACAGTGTCTGGCGGCCCCAGATGGGAAGGCTGTACTTGTGTGACACTCAAG 1922  
 QY 1757 CTGGTCTGGTGTGGGAAATCAGCCTTCTTAGCAGAGGAGGAGAAACCGTTCGAACTG 1816  
 DB 1923 CGGATCGGTGTGTGGGAAATTAAGCTTGTCTGGCTGTCTGGGGTGGCAACCGGCGACAG 1982





Db	1786	CTGGCACTCCCAAGGCATGCTGATGAGTCAAGCTGATGGTGCAGCTTCCGGACAAGA	1845
Qy	1517	TCCAGTTAGCCCTCGCCATTGATGTGAACATTTCAGCATCATCAGCAAAAGTCGACTTGTTCA	1576
Db	1846	TGGGCTGGACCTCGCCATTGACGTGAATATTCCATCGTCAGCAAAAGTGGCACTCTTCC	1905
Qy	1577	AGGTTGTGATACACAGATGATTTATGACATGTTGCTAAGATTGMAATCGTTCTCTATT	1636
Db	1906	AGGCTGTGACCGGCAGATGATCTTTGACATGCTGAAGAGGCTGCGCTCTGTAGTCTACC	1965
Qy	1637	TGCCTGCTGACTTTGCTGCAAAAAGGGAGAAATTTGCCAAGGAAATGTATATCATCAAGC	1696
Db	1966	TGCCAATGATTACGTGTGCAAGAAAGGGAGATAGSCCGGAAATGTACATCATCCAGG	2025
Qy	1697	ATGAGAAGTCCAAAGTCTTTGGAGGCCCTGATGTGACTAAAGTCTGGTTACTCTGAAG	1756
Db	2026	CGGGCAGGTGCAGGTCTCTGGTGGCCCGGATGGGAAATCCGTGCTGGTCACGCTGAAG	2085
Qy	1757	CTGGGTGGTGTGAGAAATCAGCCTCTTAGCAGCAGGAGGAGGAAACCGTCGAAC TG	1816
Db	2086	CTGATCTGTGTTGGAGAAATAAGCTTGCTGGCTGTAGGGGCGGAAACCGGCGCACAG	2145
Qy	1817	CCAATGGTGGCCCAAGGTTTGGCAATCTTTTAACTCTTAGACAAAAGACCTTCCAAG	1876
Db	2146	CTAACGTGTGGCCCATGGGTTTACCAACCTCTTCATTCTGGATGAAGAGACCTGAA TG	2205
Qy	1877	AAATTTAGTGCATTATCCAGATTTCTGAAGGATCCTCATGAAGAAAGCCAGAGTGCTT	1936
Db	2206	AAATTTCTGGTCATTAATCCCGAGTCTCAGAAGTTGCTCCGCNAGAGCCAGGCGAATGC	2265
Qy	1937	TAAAGCAGAAGGCTTAAGACCGCAGA	1961
Db	2266	TGAGAAATAACAAACAAGCCCAAGGA	2290

## RESULT 14

```

BTCNG4CGN          3253 bp      mRNA      linear      MAM 18-APR-2005
LOCUS
DEFINITION      B.taurus mRNA for CNG4c protein.
ACCESSION      X94707
VERSION      X94707.1  GI:1154636
KEYWORDS      CNG4c gene; CNG4c protein; cyclic nucleotide-gated cation channel;
              modulatory subunit.
SOURCE
ORGANISM      Bos taurus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
              Pecora; Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS      Biel,M., Zong,X., Ludwig,A., Sautter,A. and Hofmann,F.
TITLE      Molecular cloning and expression of the Modulatory subunit of the
              cyclic nucleotide-gated cation channel
JOURNAL      J Biol. Chem. 271 (11), 6349-6355 (1996)
PubMed      8626431
REFERENCE
AUTHORS      Biel,M.
TITLE      Direct Submission
JOURNAL      Submitted (08-JAN-1996) M. Biel, Inst. f. Pharmakologie und
              Toxikologie, Technische Universitaet Muenchen, 80802 Muenchen,
              Biedersteiner Strasse 29, FRG
FEATURES
source
1..3253
   /organism="Bos taurus"
   /mol_type="mRNA"
   /db_xref="taxon:9913"
   /tissue_type="testis"
   /clone_lib="cDNA in plasmid vector pCDNAII (Invitrogen)"
   /dev_stage="adult"
gene
1..3253
   /gene="CNG4c"
CDS
55..2874
   /gene="CNG4c"
   /note="modulatory subunit of cyclic nucleotide-gated

```



Qy	1037	ACAAAGCATATATCTACAGATTATTCGAACAACTGGATACCTGCTGTTTATTTCTGCACA	1096
Db	1566	GCAAAGCCTACGTTTACAGGGTTATCAGGACCACAGCCTACCTGCTCTACAGCTTACATC	1625
Qy	1097	TTAATGCCCTGTGTTTATTACTGGCTTCAAACTATGAAGAAATTGGCACTACTAGATGGG	1156
Db	1626	TGAACCTCATGTCTCTATTACTGGGCATCGGCCTATGAGGCCCTCGGCTCACTACTGGG	1685
Qy	1157	TGTATGATGGGAAGAAACGAGTATCTGAGATGTTTATTATTGGGCAGTTCGAACTTTAA	1216
Db	1686	TTTATGATGGCGTGGGAACAGTTACATTCGCTGTTTACTACTGGGCTGTGAAGACCTCA	1745
Qy	1217	TTACCAATGGTGCCCTTCGAACCAACAACCTTTATTGAAATGTTTTCAACTCTTGA	1276
Db	1746	TCACCATCGCGGCCCTCGCCGACCCCGAGCAGCTCTTTGAAATTTGCTTCCAGGGTCTAA	1805
Qy	1277	ATTTTCTTCTGGAGTTTGTGTTCTCCAGTTTAAATTTGTCAGATGAGAGATGTGATTG	1336
Db	1806	ACTATTTACGGGGCTCTTCGCTTTCTCCGTGATGATCGGACAGATGAGAGAGTAGTGG	1865
Qy	1337	GAGCAGCTACAGCAATCAGAACTACTTCGCGCCTGCTGATGATGACACCATTCGCCTACA	1396
Db	1866	GGCCGCCACCGGGGACAGACCTACTACCGCAGCTGCACTGGACAGCACCGTGAAGTATA	1925
Qy	1397	TGAACAAATTACTCCATTCTTAACTTTGTGCAAAAGCGAGTTCGGACTTGGTATGAATATA	1456
Db	1926	TGAACCTTCTACAAGATCCCGAGTCCGTCGAGAACCGGGTCAAGACCTGGTACGAATACA	1985
Qy	1457	CATGGGACTCTCAAGAACTAGATGAGTCTGATTTGCTTAAAGACCTACCAACTACGG	1516
Db	1986	CCTGGCACTCCCAAGGATGCTGATGAGTCAAGCTGATGGTGCAGCTTCCGGACAAGA	2045
Qy	1517	TCAGATTAGCCCTCGCCATTGATGTGAACCTTCAGCATCATCAGCAAGTCGACTTGTTC	1576
Db	2046	TGCGCTGGACCTCGCCATTGACGTGAACCTATTCATCGTCAGCAAGTGGCACTCTTCC	2105
Qy	1577	AGGGTTGTGATACACAGATATTATGACATGTTGCTTAAGATTGAATCCGTTCTCTATT	1636
Db	2106	AGGGCTGTGACCGGCAGATGATCTTTGACATGCTGAAGAGGCTGCGCTCTGTAGTCTACC	2165
Qy	1637	TGCTGTGTGACTTTGTCTGCAAAAGGGAGAAATGGCAAGGAAATGTATATCATCAAGC	1696
Db	2166	TGCCCCAATGATTACGTGTGCAAGAGGGGAGATAGGCCGGGAATGTATCATTCAGG	2225
Qy	1697	ATGGAGAAGTCCAAAGTTCTTTGGAGCCCTGATGTACTTAAAGTTCTGGTTACTCTGAAAG	1756
Db	2226	CGGGCAGGTGCAGTCTTGGGTGGCCCGATGGGAATCCGTGCTGGTGAAGTGAAGG	2285
Qy	1757	CTGGTCTGGTGTGGAGAAATCAGCCTTCTAGCAGCAGGAGGAGGAACCGTGCAGCTG	1816
Db	2286	CTGGATCTGTGTTTGGGAGAAATAAGCTTGTGGCTGTAGGGGGCGGAACCGGCGCACAG	2345
Qy	1817	CCAATGTGTGGCCACGGGTTTGCCAACTCTTTTAACTCTAGACAAAAGACCTCCCAAG	1876
Db	2346	CTAACGTGGTGGCCCATGGGTTTACCAACCTCTTTCATCTGGATGAAGGACCTGAATG	2405
Qy	1877	AAATTCCTAGTGCAATTATCCAGATTCTGAAGGATCTCTCATGAAGAAAGCCAGAGTGTCTT	1936
Db	2406	AAATTCCTGGTGCAATTATCCCGAGTCTCAGAAAGTTGCTCCGCAAGAGGCCAGGCAATGC	2465
Qy	1937	TAAAGCAGAGGCTTAAGACCGCAGA	1961
Db	2466	TGAGAAATAACACAAGCCCAAGGA	2490

**THIS PAGE BLANK (USPTO)**